

Controlling the narrative

Dr Norman Lewis

The EU's attack on online speech

May 2024



Controlling the narrative: The EU's attack on online speech

Dr Norman Lewis

Contents

1 Executive summary	7
2 Introduction	11
2.1 Disinformation: in the name of protecting children	13
2.2 The structure of the report	16
3 The hate speech narrative	20
3.1 The rise and rise of hate speech	20
3.2 What is hate speech?	23
4 The history of the EU's turn against freedom of speech	27
4.1 The Universal Declaration of Human Rights (UDHR) and free speech	28
4.2 The International Covenant on Civil and Political Rights (ICCPR)	30
4.3 Hate Speech Laws and the Rule of Law	31
4.4 The Internet and Brexit	34
5 The EU's "operating system" for controlling speech	36
5.1 The EU's code of misconduct – the invisible hand of outsourced technocratic censorship	36
5.2 The EU 'Ministry of Truth' – the unholy alliance of unaccountables and the privatisation of state censorship	39
5.3 The case of Sanchez v France (2021)	44
5.4 The case of Beizaras and Levickas v Lithuania (2020)	47
5.5 Big Brother and Big Tech	49
5.6 The institutionalisation of freedom from speech	52
6 AI and the automation of language policing	55
6.1 A true disinformation narrative	56
6.2 AI and automating authoritarianism	60
6.3 Brussels on the warpath	65
7 The forthcoming elections and the battle of the narratives	70
7.1 The European Media Freedom Act (EMFA)	72
7.2 The EU's hate campaign against the peoples of Europe	75
7.3 So, what is to be done?	76
8 Endnotes	78
9 About the author	82

1 Executive summary

The issue of free speech has always been a contest about who decides what can be said, heard or thought in society. The European Union's focus on curbing what it calls 'hate speech' and 'disinformation' is the latest form of this struggle. Under the guise of upholding civilised norms of behaviour, the EU is institutionalising laws against hate speech and disinformation which represent a fundamental attack on free speech and democracy in Europe.

A package of laws, regulations and agreements between EU institutions and Big Tech represent an attempt by EU elites to determine what Europe's 448 million people can or cannot say online. Further sweeping regulations on online speech are planned. The justification they give is the need to protect European democracy from hate speech and misinformation. But behind these invocations of democracy in fact lies a profoundly anti-Democratic attitude towards European citizens.

Rather than Europe being under attack from "hate speech", European citizens are under attack from the hateful attitude of EU elites. The powers-that-be look down on European citizens as infants easily susceptible to manipulation who need to be insulated from harmful speech and ideas. This report aims to challenge the Brussels 'hate speech' narrative.

The policing of speech to attempt to socially engineer political outcomes has become the modus operandi of the EU's fragile technocratic oligarchy, who fear any open and unpredictable debate that may raise fundamental questions about their right to rule and the legitimacy of Brussels' policies on key issues from the green deal to mass migration. This fear has become

heightened in the run-up to the June elections to the European Parliament, which are predicted to see a surge in support for national parties opposed to centralised EU control.

This challenge to the ruling EU orthodoxy has led to demands for ever-more intervention in European debate. This is why the censorship operating system – the panoply of laws, unaccountable NGOs and Big Tech – outlined in this report is only set to expand. The censorious crusade against free speech is not a temporary phenomenon but is at the core of how the EU and its institutions now operate.

The report has four key points:

- First, the hate speech narrative is not about good manners or a system of government that elevates civilised behaviour to protect citizens. It is a politically motivated crusade to institutionalise an EU ‘Ministry of Truth’ whose goal is to protect the EU and its central institutions from free speech.
- Second, since the EU came into existence, the evolution of hate speech laws has been driven by anti-democratic impulses. The EU elite are perpetually fearful of the views and opinions of European citizens. Since the end of the Second World War, European elites have seen their mission as protecting Europe from the “dangers” of untrammelled democracy. Brussels has thus become institutionally afraid of the open-ended unpredictability of free speech and elections. This has only intensified in recent years as, across the EU, political forces are on the rise that view European culture and history differently and question the status quo.

- Third, this censorious dynamic can only increase in the future as it becomes automated and automatic. This crusade without an end is about to be boosted by the automation of hate speech detection online, through the application of Artificial Intelligence. Curiously, this is one area where the EU's default risk-aversion and precautionary approach to innovation does not apply. Weaponising AI to advance the policing of speech represents a real and present danger to the future of European democracy.
- Fourth, the battle with the Eurocrats over the narrative about hate speech and disinformation is one that we cannot afford to lose. It is a battle that has to be won by those who understand how central free speech remains to democratic rights and freedom. More speech, not freedom from speech, is our best defence not only against hateful speech but against an increasingly authoritarian EU oligarchy which is happy to sacrifice free speech and democracy if it leaves the status quo intact.

As the report concludes, the stakes are very high. The malicious and hateful prejudice of the EU elite that ordinary people are too ignorant, stupid and prone to easy manipulation by demagogues needs to be forcefully countered.

During the coming elections, the goal should be to expose every attempt to muzzle views and speech deemed out of order by Brussels and their Big Tech minions.

By spreading its disinformation narrative, the Brussels elite can itself be accused of propagating 'disinformation' or 'fake news'. The real threat to the EU elections and the future of European democracy is the EU's crusade against hate speech and disinformation. The real issue is who controls what can or cannot be said or thought in Europe.

EXECUTIVE SUMMARY

The best defence for democracy is always free speech. Rather than those who wish for less speech or controlled speech, we advocate for more speech and freer speech. More speech conducted openly in the court of public opinion is the only long-term foundation for protecting democracy in Europe.

2 Introduction

Towards the end of January, Josep Borrell, the EU High Representative of the Union for Foreign Affairs and Security Policy/Vice-President of the Commission, made an important speech in Brussels on the dangers of disinformation and foreign interference in the forthcoming elections.¹ For Borrell, ‘Foreign Information Manipulation and Interference’ (FIMI) represents ‘one of the most significant threats of our time ... because these actors actively seek to undermine democracy, proposing an alternative model, making people mistrust everything – mistrust institutions and toxically infiltrating our societies, to poison them’. According to Borrell, disinformation is ‘not about a bomb that can kill you; it is about a poison that can colonise your mind’. Malicious content circulated at the speed of light now ‘spreads like cancer and puts the health of our democracies at risk’ he asserted, from which he concluded that the EU is now in a ‘battle of narratives, and this battle has to be won’.

There is no question that a ‘battle of narratives’ is taking place in Europe, which will intensify in the forthcoming elections and beyond. The threat of generative AI, and the manipulation of information Borrell fears, echoes a global narrative of panic on the part of those in control of society. If we were to believe the wisdom of the 1,500 experts surveyed in the World Economic Forum’s 2024 Global Risks Report,² for example, AI-powered lies and manipulation constitute the gravest threat to humanity. Apparently, generative AI will ‘trigger the next misinformation nightmare’,³ that people ‘will not be able to know what is true anymore’,⁴ and that we are facing a ‘tech-enabled

Armageddon'.⁵ In the hands of malicious powers like Russia or China, their ability to spread hate and disinformation – ‘poison that can colonise’ the minds of ordinary Europeans – is what is keeping Borrell and his fellow Commissioners awake at night.

The announcement by Commission Vice President Věra Jourová to launch a ‘democracy tour’ across Europe is part of this narrative. Linking Russian ‘disinformation’ with Germany’s far-right Alternative for Germany (AfD), she suggested that unless this was countered, ‘democratic processes might give the blessing to the end of democracy.’⁶ Like the EU’s Authority for European Political Parties and European Political Foundations, freedom of speech is now considered dangerous because it ‘can be weaponised by injecting and fuelling false information’.⁷

The irony of defending democracy by curbing free speech seems to elude the Brussels elite. But this is not an oversight. It is Orwellian newspeak. It is the essence of a disinformation narrative pursued by the EU oligarchy for the forthcoming elections and beyond.

Under the guise of the danger of FIMI, the real target of the EU elite’s narrative, which is at home, is hidden from view. It is the ordinary voters

The actual ‘poison’ threatening the forthcoming elections and the future of European democracy is the EU’s crusade against hate speech

of Europe’s member states who they assume lack the intellectual or moral independence to critically evaluate the views they are exposed to and thus can easily be swayed to think or vote incorrectly. Linking

foreign intervention to the rise of ‘far right’ populism is a disingenuous smokescreen that hides their hateful antipathy towards ordinary Europeans, whom they look down upon as infants susceptible to manipulation who need to be protected from themselves by insulating them from harmful speech

and ideas. It is not the manipulative ways of Putin that is the most significant danger in Europe. It is the self-appointed guiding hand of Big Brussels and their experts who think they know what's best for Europe and who are willing to undermine democracy itself through criminalising speech because that, after all, is the medicine needed to keep the children on the path of truth and federal righteousness.

The actual 'poison' threatening the forthcoming elections and the future of European democracy is the EU's crusade against hate speech dressed up in Borrell's AI-fearing narrative. The real issue is not disinformation but who controls it and the ends that serves.

If proof is needed of this contempt, indeed hatred, towards ordinary Europeans, then the recent case of Commissioner for Home Affairs Ylva Johansson's disinformation campaign on X in support of passing her beleaguered Child Sexual Abuse Regulation (CSAR)⁸ demonstrates this all too clearly.

2.1 Disinformation: in the name of protecting children

The problem Johansson faced was the pushback her proposed CSAR legislation had caused. The CSAR aimed to curb the dissemination of child pornography by automatically monitoring the digital correspondence of all European citizens with the use of AI technology. At its heart were controversial chat controls.⁹ These aimed to infringe upon end-to-end encryption on digital communication apps, such as WhatsApp, iMessage, Instagram, TikTok and X, by automatically scanning the digital communications of all EU citizens to detect child sexual abuse. Given what this attack on privacy represented, the backlash was understandable and unprecedented.¹⁰ Even 80 NGOs signed an open letter opposing the CSAR.¹¹

Minutes acquired from the Council of the European Union meeting on 14 September revealed insufficient support for the CSAR, which meant it would fail to pass.¹² This galvanised Johansson and the Commission to act.

That the Commission was willing to violate its flagship legislation reveals how pragmatic they are about the actual applicability of the ‘rule of law’ when it suits them

A day after the Council meeting, Johansson commissioned a paid advertising campaign on X targeting the citizens of the Netherlands, Sweden, Belgium, Finland, Slovenia, Portugal and the Czech Republic – all the countries that the minutes showed were unwilling to vote for the proposed legislation. The online campaign, produced in multiple languages

and viewed more than four million times, used emotionally charged images of children juxtaposed with adults who appeared to be predators while insisting that ‘time is running out’ backed up by a clock ticking. The targeting of the ad campaign suggested that opponents of the proposed legislation did not want to protect children. The aim was to pressurise the national governments that had sought reforms of the proposals during that European Council meeting on September 14 to change tack.

A survey¹³ conducted by a company called Savanta¹⁴ in support of this campaign was manipulated to create the impression that most European citizens backed the proposals. This was deliberate disinformation. Alternative surveys by Novus¹⁵ showed the opposite: there was virtually no support for the proposal among the European population.

However, what was even more revealing was how Johansson and the European Commission used unlawful ‘microtargeting’ to ensure that the ads did not appear to people who care about privacy. People targeted with this paid-for ad would *not* be people interested in ‘Julian Assange’ or labelled as ‘Eurosceptics’: those who had expressed an interest online in ‘Next’,

‘Brexit’ and ‘Spanexit’ or in Victor Orbán, Nigel Farage, or the German political party AfD.

In short, the campaign deliberately set out to manipulate ordinary people, whom the Commission contemptuously felt could be hoodwinked into acting as stooges for Brussels.

But it also demonstrated another critical truth. These actions violated the 2024 Digital Services Act (DSA)¹⁶, which enables Brussels to dictate what can or cannot be said online with huge fines for social media companies if they fail to comply, and the General Data Protection Regulation (GDPR)¹⁷ which, among many things aims to protect privacy in Europe. The Vienna-based European Center for Digital Rights (None Of Your Business NYOB) – led by lawyer turned activist Max Schrems – filed a complaint with the European Data Protection Supervisor (EDPS) because the microtargeting of social media users on the basis of their religious beliefs or political views is specifically outlawed under the GDPR.¹⁸ Mickey Manakas, spokesperson for NYOB, reiterated in an interview with the European Conservative the basic point that ‘the use of sensitive data, such as political opinions and religious beliefs’ which are ‘specifically protected by the EU GDPR ... can’t legally be used for political micro-targeting’.¹⁹ The fact that the Commission was willing to violate its flagship legislation reveals not just cynicism but also how pragmatic they are about the actual applicability of the ‘rule of law’ when it suits them.

This demonstrates that what matters for the EU elites is the engineering of their desired policy outcomes – against the will or concerns of their citizens. Controlling the narrative to socially engineer the status quo is what keeps the EU elite awake at night.

The hate speech and disinformation narratives are just that – narratives. They aim to control what can or cannot be said in the European political

arena to ensure that the Brussels federal and cosmopolitan view of the world dominates.

This report aims to offer a critical analysis of the crusade against hate speech and disinformation to provide a much-needed counter-narrative to Brussels' anti-democratic authoritarianism. It is a deceptive narrative that pretends to uphold democracy but represents an undeclared war on the foundations of democracy – on free speech and member states' demos who cannot be trusted to think and act in accordance with the outlook of their overlords in Brussels.

2.2 The structure of the report

- In Chapter 3, we interrogate the concept of hate speech. The remarkable thing about hate speech is the absence of an agreed definition of what it is. This is not an oversight but a deliberate outcome. The real purpose of hate speech laws is not curbing 'hateful' speech but free speech. It is not aimed at the protection of minorities but at controlling the speech of the majority. The lack of precision and its ever-widening shopping list of vulnerabilities and protected characteristics is determined by the overriding objective of silencing controversial views, shutting down unpredictable debate, and not containing hate.
- Chapter 4 charts the evolution of hate speech laws in the EU and shows how freedom from speech displaced any free speech defence in one generation. This evolution also demonstrates how the EU's right to determine what could or could not be said always trumped due process. Hate speech laws were always a means to an end rather than an end. The real goal was the systematic attempt to ensure the EU and its institutions would determine what could and could not be said in Europe.

- Chapter 5 examines how the EU has developed an extensive system for controlling speech – online and off. Most noteworthy in this system is an “unholy alliance” between EU institutions, Big Tech, and NGOs. The EU sets the tone for the control of speech online, which is implemented by Big Tech platforms (sometimes under threat of EU legislation and sometimes in direct response to it). Both parties are pressured by a network of NGOs, who have in fact become embedded in the system of censorship itself. In all three cases, what can be said online is thoroughly controlled by completely unaccountable bodies. Such a wide-ranging system of speech control, completely removed from any democratic oversight, is unprecedented in Europe’s post-war history.
- Chapter 6 examines how this inner censorship dynamic manifests itself in the wholesale embracing of hate speech detection and takedown through the application of Artificial Intelligence wielded by unaccountable social media platforms. While the EU is characterised by its risk-averse embrace of the precautionary principle, all the fears expressed about the dangers of AI, as noted above, disappear when it comes to wielding this to serve the EU elite’s goal of controlling what can or cannot be said online. The only risk involved for them is the risk of not clamping down on free speech. Policing language – censorship – as an attempt to socially engineer outcomes has become the modus operandi of the EU. Freedom of speech has been displaced by the automation and institutionalisation of freedom from speech.
- In Chapter 7, we return to the battle of the narratives Borrell initiated in the run-up to the forthcoming elections. The threat of malicious foreign players intervening in the internal affairs of Brussels is just a ruse: the EU bureaucracy has no problem with foreign involvement in its internal affairs,

for example, through foreign-backed, financed NGOs who support their crusade against hate speech. Its European Media Freedom Act (EMFA) is an Orwellian expression of the fact that the last thing they want is media pluralism. Their quest is to concentrate power in Brussels. The only

We have seen the creation of an authoritarian, anti-democratic system of government which is consciously hidden from view by disinformation narratives sourced in Brussels

freedom Brussels will countenance is their freedom to determine the narrative, to dictate what can and cannot be said and thus to protect their freedom to impose their technocratic values from above.

To achieve this, the EU has erected

a legalised ‘censorship operating system’ enforced through an unholy alliance of unaccountables – from the EU Commission through its network of unelected experts and NGOs to equally unaccountable Big Tech and EU Courts.

The report concludes that the attack on free speech is not episodic but systemic. Its ever-expanding scope is driven by a deep insecurity on behalf of the EU ruling elite. The attempt in January by the European Parliament to extend the list of EU-wide crimes to include all forms of ‘hate crime’ and hate speech as a means of forcing member states to comply with Brussels’s censorship diktat demonstrates how inexorably the EU’s self-expanding censorship dynamic operates in practice.

We have seen the creation of an authoritarian, anti-democratic system of government which is consciously hidden from view by disinformation narratives sourced in Brussels. Even before the elections started, the Commission announced an investigation under the aegis of the Digital Services Act (DSA) into X for allegedly breaking EU law on disinformation.

Formal proceedings began against TikTok to assess whether there had been any similar breach of the DSA.²⁰

Today, all member states currently criminalise hate speech on the grounds of race, colour, religion, descent, national or ethnic origin. Only 20 of them explicitly include sexual orientation in hate speech legislation, while 12 include gender identity and a mere two cover sex characteristics. Their impact on public political debate has been chilling. In January this year, for example, in Albstadt, Germany, a 50-year-old man was placed on probation after being sentenced to three months in prison because he called asylum seekers “freeloaders” on X in November 2022.²¹ In another case, in Hamburg, an anonymous user of Twitter who posted ‘Du bist so 1 Pimmel’ (“You are such 1 dick’) about Andy Grote, the Minister of the Interior, for his hypocrisy, resulted in a dawn police raid at the suspects home to root out this hate speech.²² The case was finally laughed out of the courts. But the battle over the narrative of hate speech is no laughing matter.

It is the court of public opinion where truth, facts, and ideas can only be genuinely contested

Borrell is right about one thing: this is, indeed, a ‘battle of narratives, and this battle has to be won’ – not by Eurocrats – but by democrats who value and understand how central free speech remains to democratic rights and freedom. It is the court of public opinion where truth, facts, and ideas should and can only be genuinely contested.

This report places trust in ordinary people’s ability to weigh up different sides of an argument and make independent decisions about where their interests lie. More speech, not freedom from speech, is the only defence against an increasingly authoritarian, anti-democratic EU oligarchy which is happy to sacrifice free speech and democracy if it leaves the status quo intact. This is the narrative that needs defeating.

3 The hate speech narrative

In this chapter, we examine the concept of hate speech and the criminalisation of certain types of speech. The most striking thing about hate speech as a concept is its terminological imprecision. A concept that should be clear and should define with legal precision what forms of content should be regulated, turns out to be unclear and imprecise. This curiosity raises questions as to what purpose such legislation serves. This is where the narrative on hate speech becomes important.

3.1 The rise and rise of hate speech

In recent years, the discussion about hate speech has proliferated. Today, it is impossible to get away from the idea of hate speech in the media or academia. In recent years, references to hate speech, case law, and incidents have dramatically expanded in academia, the media and online. For example, a search for hate speech articles on the Semantic Scholar website – the free, AI-powered research tool for scientific literature – returns 31,400 papers already for 2024 that deal in some shape or form with the topic. Over the last ten years, the site has recorded 330,000 papers.²³ A similar search from 2020 onwards yielded 24,500 results for books and journal articles on Google Scholar. Outside of academia, search results on Amazon UK alone for books with hate speech in their title returned 862 results. The *Guardian* newspaper returns 194 articles this year alone with ‘hate speech’ in their titles.²⁴ *Politico*, across all its publications over a more extended period, records 1,766 articles.²⁵

Much of this media coverage relates to the reporting of incidents of hate speech crimes. In the wake of legislative changes, these statistics have grown equally dramatically. For example, after the implementation of the German Hate Speech Law (Netzwerkdurchsetzungsgesetz) in 2020, which will now be supplanted by the EU's DSA as mentioned above, in the State of Bavaria alone, 2,435 official investigations took place by district attorneys. This resulted in 488 formal criminal charges and 324 convictions by the court.²⁶ In 2023, the State of Bavaria installed an online platform where citizens can report cases of hate speech: 10,619 reports were forwarded to the police, which resulted in a 75 per cent increase in criminal charges from 2022.²⁷

The European-wide agency, the European Union Agency for Fundamental Rights (FRA), founded in 2007 and headquartered in Vienna²⁸ (whose role, among others, is to 'raise rights awareness at the EU, national

The EU, and its associated NGOs, want to suggest that hate speech is an under-reported but massively growing problem

and local level'²⁹), argues that 'millions of people across the EU experience hate-motivated violence and harassment' which can be due to their 'ethnic or immigration background, skin colour, religion, gender, sexual orientation or disability'.³⁰ Their

recent 'Fundamental Rights Survey'³¹ asserts that some 'minority groups experience twice as much violence as people generally', while 'up to 9 in 10 people do not report being attacked', which suggests that hate speech and related crimes are not only enormous but remain hidden from public view.

The EU, and its associated NGOs, want to suggest that hate speech is an under-reported but massively growing problem in society. The idea is that doing something to address this hate pandemic is necessary and desirable.

For many upholders of the criminalisation of hate speech, hate speech is presented as a slippery slope which inevitably ends up in discriminatory

practices and violence. A statement by the European Commission Against Racism and Intolerance (ECRI)³², (a Non-Governmental Organisation [NGO] we will discuss in more detail in chapter 5), at its 85th plenary meeting in March 2021, is paradigmatic in linking hate speech and violence:

*Failure to prevent and combat ultra-nationalistic and racist hate speech and violence in this context further fuels hatred between communities, leading to discriminations on the grounds of national, ethnic, linguistic or religious background or of citizenship, which can entail other grave violations of the European Convention on Human Rights and is likely to result in retaliatory practices and the descent into an inexorable vortex of violence and tragedies.*³³

This is strong stuff. But the asserted link between hate speech and an ‘inexorable vortex of violence and tragedies’ is somewhat undermined by

The asserted link between hate speech and hate crime is just that, an assertion

the earlier qualification that this is ‘likely’, not definite. In a similar assertion, ECRI elsewhere argues that while hate speech ‘poses grave dangers for the cohesion

of a democratic society’, if left unaddressed, this ‘can lead to acts of violence and conflict on a wider scale’.³⁴ Yet, no evidence is presented to show this outcome or its inexorability.

The asserted link between hate speech and hate crime is just that, an assertion. But it is now accepted as an uncontested fact. The European Commission, when urging the European Parliament to consider making hate speech an EU-wide crime, argued as follows:

Hate speech can lead not only to conflict, but also to hate crimes. Evidence points to a ‘pyramid of hate’ or a ‘ladder of harm’ starting from acts of bias (eg bullying, ridicule, de-humanisation) and discrimination (eg economic, political), moving up towards bias

*motivated violence, such as murder, rape, assault, terrorism, violent extremism, even genocide.*³⁵

As Frederick Attenborough points out, this reference to a ‘pyramid of hate’ is taken from U.S. social psychologist Gordon Allport’s ‘Scale of Prejudice,’ first developed in 1954,³⁶ when free speech was no longer seen as an unambiguous virtue. The advocacy of therapeutic censorship which Allport promoted was based on his influential study *The Nature of Prejudice* (1954), where he warned, ‘then even a relatively mild verbal attack may start an unimpeded progression towards violence’.³⁷ His conclusion that ‘any program for the reduction of prejudice must include a large measure of semantic therapy’ anticipated today’s drive to police language as a necessary precautionary step to halt the ‘inexorable’ slide to violence and social conflict. The pathologisation of speech will be examined in more detail in chapter 6.

The assertion of the link between speech and violence lacks evidence to support it. But things become even less clear when we try to define what hate speech is. Instead of the precision needed to give these assertions some credibility, we find contradictory and shifting definitions.

3.2 What is hate speech?

A factsheet produced by the European Court of Human Rights (ECtHR) – the body that has heard hundreds of hate speech and freedom of expression cases across Europe – concedes that, despite laws being in place in Europe for decades, there remains ‘no universally accepted definition of the expression hate speech’.³⁸ In 2015, UNESCO published a manual addressing online hate speech, acknowledging that ‘the possibility of reaching a universally shared definition seems unlikely’.³⁹ Indeed, an earlier fact sheet of the ECtHR made the point that identifying hate speech can be difficult because some kinds of speech do not necessarily manifest themselves ‘through the expression

of hatred or of emotions’. The paper also conceded that hate speech can be concealed in statements that ‘may seem to be rational or normal’.⁴⁰

The FRA mentioned above has tried to inject some precision without much success. It has stated that ‘hate speech’ refers to ‘the incitement and encouragement of hatred, discrimination or hostility towards an individual that is motivated by prejudice against that person because of a particular

This is an arbitrary broad *ladder of harm* far removed from legal clarity or unambiguous objective criteria

characteristic’.⁴¹ In another document, FRA states that it ‘includes a broader spectrum of verbal acts ... [including]

disrespectful public discourse’⁴² while another paper argues that ‘there is currently no adequate EU binding instrument aimed at effectively countering expression of negative opinions.’⁴³ But the Oscar for the greatest hate speech tautology must go to the Organization for Security and Co-operation in Europe, whose report ‘Hate Crime Victims in the Criminal Justice System’⁴⁴ clarifies things as follows:

Hate crime victims are individuals who have suffered harm as a result of a hate crime and consequently require protection, specialist support and the opportunity to receive compensation.

One definition that appears more coherent is the one provided by ECRI mentioned above. Its ‘General Policy Recommendation No. 15 on Combatting Hate Speech’ defines hate speech as follows:

The advocacy, promotion or incitement, in any form, of the denigration, hatred or vilification of a person or group of persons, as well as any harassment, insult, negative stereotyping, stigmatization or threat in respect of such a person or group of persons and the justification of all the preceding types of expression, on the grounds of “race”, colour, descent, national or ethnic origin, age, disability, language, religion

*or belief, sex, gender, gender identity, sexual orientation and other personal characteristics or status.*⁴⁵

What precision exists to determine what constitutes ‘denigration’, ‘vilification’, ‘insult’, ‘negative stereotyping’ or ‘stigmatization’ in this shopping list of characteristics needing protection? This is an arbitrary broad ‘ladder of harm’ far removed from legal clarity or unambiguous objective criteria applicable in all circumstances.

In his book *Censored*, Paul Coleman, the senior counsel and deputy director of the Alliance Defending Freedom, who specialises in international human rights and European law, makes the following insightful point about the broad and sweeping nature of hate speech captured above. He states that hate speech:

*... does not necessarily manifest itself through the expression of hatred, and it may appear rational and normal; it is always motivated by hatred, providing that the hatred is targeted at groups of the state’s choosing; and although it is impossible to define, “hate speech” may include denigration, disrespect, vilification, negative opinion, glorification, denial, trivialization, justification, condonation, incitement, discrimination, hatred, hostility, and insult.*⁴⁶

The critical insight is his reference to ‘the state’s choosing’. The real purpose of ill-defined hate speech laws is not legal protection but criminalising speech

The real purpose of ill-defined hate speech laws is not legal protection but criminalising speech the authorities deem problematic

the authorities deem problematic.

Hate speech, it turns out, is a concept in the eye of the beholder. The narrative that purports to reflect a sensitivity to minority groups and a new decent inclusionary ethic has little to do with

that. It masks the fact that the logic of hate speech laws is to ensure the right

of the state authorities to determine what type of speech is acceptable or should be criminalised.

Hate speech laws are characterised by imprecision: their real purpose is not curbing ‘hateful’ speech but *free* speech

From this perspective, it becomes clear why hate speech laws are characterised by imprecision. The real purpose of hate speech laws is not curbing ‘hateful’ speech but *free* speech. It is not aimed at the protection of minorities but at controlling the speech of the majority. The lack of precision and its ever-widening shopping list of vulnerabilities and protected characteristics is determined by the overriding objective of silencing controversial views, shutting down unpredictable debate, and not containing hate.

Hate speech laws are the legal form through which the EU bureaucracy attempts to police the speech of the European people. It is based on an elitist self-belief that they and their experts know what’s best for society and thus have the right and duty to decide what is acceptable to ensure social harmony. Hate speech laws are a fundamental part of the EU’s armoury to safeguard their right to determine what can or can’t be said in Europe. The only ‘inexorable vortex’ of hate speech laws is the need for the EU elite to decide what narrative rules in Europe.

The historical evolution of hate speech laws in the EU demonstrates this reality. This is the focus of the following chapter. We will then be able to explain what drives this and why the shift from free speech to today’s freedom from speech occurred. This, in turn, will explain why an EU ‘Ministry of Truth’ has emerged, which is systemic, not nice to have, but a condition for the EU’s future survival.

4 The history of the EU's turn against freedom of speech

In this chapter, we chart the evolution of hate speech laws in the EU and show how freedom from speech displaced any free speech defence in one generation. This evolution also demonstrates how the EU's right to determine what could or could not be said always trumped due process. Hate speech laws were always a means to an end rather than an end. The real goal was the systematic attempt to ensure the EU and its institutions would determine what could and could not be said in Europe.

While limitations on speech have existed for centuries, the limitation of 'hateful speech' is a modern post-Second World War phenomenon.

In the aftermath of World War II and the Holocaust, the need to address

While limitations on speech have existed for centuries, the limitation of 'hateful speech' is a modern post-Second World War phenomenon

the atrocities committed during the war was pressing. The idea that certain speech should be outlawed arose in the debate about the future protection of human rights in the post-war world. The discussion that unfolded and culminated initially in establishing the Universal Declaration of Human Rights

(UDHR) in 1948 set the tone and content which would eventually result in the internationalisation of hate speech laws.⁴⁷

While this process was undoubtedly boosted with the emergence of the Internet and social media towards the end of the 20th Century, this simply sped up the existing process of weaponising linguistic policing rather than causing it.

It is factually correct to brand hate speech legislation as a ‘tenacious Soviet legacy’⁴⁸ – arising as it had in the Soviet experience of Nazi barbarism. However, the historical irony is that it is the West today, and the EU in particular, who, in their bid to justify their desire to protect society from hate speech, revert to some of the same authoritarian, anti-democratic arguments the West vigorously opposed back then.

4.1 The Universal Declaration of Human Rights (UDHR) and free speech

The core debate of the UDHR revolved around the question of tolerance.⁴⁹ The tension between freedom of expression and limiting speech became apparent during the discussion over Article 19 (freedom of expression) and Article 7 (protection against discrimination). Every attempt by the Soviets

The framers of the UDHR were not prepared to single out speech labelled loosely as ‘fascist’ as being unworthy of protection

to introduce amendments intended to deny freedom of speech and assembly to those labelled as ‘fascists’ was defeated. As Coleman points out, and which is so starkly removed from today’s perspective, the majority equally virulently

anti-fascist⁵⁰ view held that ‘tolerance should mean tolerating even the intolerant’.⁵¹

The framers of the UDHR were not prepared to single out speech labelled loosely as ‘fascist’ as being unworthy of protection. Every attempt by the Soviets to link free speech and freedom of association with ‘fascism’, which should include punitive sanctions against dangerous ideas and their diffusion, was rejected.⁵² The final version of Article 19 did not explicitly exclude any particular people or group from protection.

However, the Soviets more successfully drafted Article 7 in the Sub-Commission on the Prevention of Discrimination and the Protection of Minorities. Their success lay in introducing anti-discrimination language into the UDHR. An amendment introduced by Alexander Bogomolov, the Soviet Member of the Commission on Human Rights, explicitly coupled free speech and hate speech. In pathological language that anticipates present-day approaches, Bogomolov argued that

*Between Hitlerian racial propaganda and any other propaganda designed to stir up racial, national or religious hatred and incitement to war, there was but a short step. Freedom of the press and free speech could not serve as a pretext for propagating views which poisoned public opinion.*⁵³

Although this proposal was rejected in this form, a compromise amendment was adopted, and the words ‘and against any incitement to such discrimination’ were added to Article 7. Thus, while the final version of Article 19 contained no clauses limiting free speech, Article 7 ultimately included a right to be protected from ‘incitement to discrimination.’

The assumption of a link between free speech and the propagation of hateful views which ‘poisoned public opinion’ was now accepted as a legal principle

The assumption of a link between free speech and the propagation of hateful views, which ‘poisoned public opinion’, was now accepted as a legal principle. This link in the

unfettered defence of free speech was to prove to be the fatal gap through which subsequent hate speech laws would flow. Indeed, Holocaust denial, which was instituted in the EU after 1995 – the first systematic expression of hate speech laws – is very much based on this assumption.

4.2 The International Covenant on Civil and Political Rights (ICCPR)

The passage of The International Covenant on Civil and Political Rights (ICCPR) in 1966, which came into force a decade later, is an important milestone. Unlike the UDHR, the passing of the ICCPR was binding, which meant that any state ratifying these treaties was required to take positive measures to introduce hate speech laws into their national criminal law. What was different about the final version is that the European communist nations were able to generate enough support to pass amendments prohibiting hate speech while those standing out for free speech were in decline.

As Coleman recounts, the representative from Yugoslavia argued, presciently anticipating today's debate, that while incitement to violence should be prohibited, 'it was just as important to suppress manifestations of hatred which, even without leading to violence, constituted a degradation of human dignity and a violation of human rights.'⁵⁴ This argument about protecting what is loosely termed 'human dignity' forms the bedrock of the imprecision in contemporary hate speech laws today, as discussed in the previous chapter.

At the time, however, defenders of free speech, like Eleanor Roosevelt, fervently cautioned against such imprecision and added that 'any criticism of public or religious authorities might all too easily be described as incitement to hatred and consequently prohibited.' Not only was this 'unnecessary, it was

It is noticeable that already, by this time, those arguing for free speech were in decline

also harmful.'⁵⁵ But the summary of the General Assembly report rejected this. It argued instead that any 'propaganda which might incite discrimination or

hostility would likely incite violence and should therefore be prohibited.'⁵⁶

The outcome was Article 20(2), which states that 'any advocacy of national,

racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law’.

It is noticeable that already, by this time, those arguing for free speech were in decline. Lady Gaitskell’s opposition to Article 4 of the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD), passed in 1965 by the UN,⁵⁷ shows this clearly. Her argument that allowing the state to use coercive means to eradicate speech that is deemed

Liberal Democracies that once opposed the attack this represented to freedom of speech would now become the most fervent advocates of an ever-expanding censorious dynamic

by the state to be hateful ‘infringed the fundamental right of freedom of speech’ went unheeded. Instead, Article 4 of the ICERD and Article 20(2) of the

ICCPR established in law the empowering of the state to censor speech by criminalising it. This marked the start of the internationalisation of the criminalisation of hate speech.

Thus, in less than three decades, the defence of free speech had slipped dramatically. But now, driven by new political needs in Europe, not only would hate speech laws expand, but the threshold of what constitutes illegality would consistently diminish. Liberal Democracies that once opposed the attack this represented to freedom of speech would now become the most fervent advocates of an ever-expanding censorious dynamic.

4.3 Hate Speech Laws and the Rule of Law

A new dynamic came into play when the EU, as it is now constituted, came into being through the 1993 Maastricht Treaty. The top-down political union necessitated the consolidation of a new Federal legal system in which hate speech laws were to play a vital role. A fundamental change occurred with

the passing of the 2008 Framework Decision on Racism and Xenophobia, which now had to be incorporated into each member state's criminal law.⁵⁸

This framework set out that the scope of hate speech encompassed:

*combating certain forms and expressions of racism and xenophobia ... [and] all conduct publicly inciting to violence or hatred directed against a group of persons or a member of such a group defined by reference to race, colour, religion, descent or national or ethnic origin.*⁵⁹

While the Framework Decision still upheld relatively high thresholds of what constituted criminality – conduct needed to be intentional – member states were also given the widespread and imprecise choice to punish conduct that could be interpreted as disturbing public order or is ‘threatening, abusive, or insulting’.⁶⁰ As discussed above, this loose formulation would characterise hate speech legislation from then onwards.

However, one of the most significant things about the Framework was that it became part of the EU's legal system. It was now obligatory for EU member states to criminalise certain forms of hate speech and hate crime, making incitement to racism and xenophobia punishable by law.⁶¹ There was

It was now obligatory for EU member states to criminalise certain forms of hate speech and hate crime, making incitement to racism and xenophobia punishable by law

no leeway or space for interpretation when implementing such legislative changes at the member state level. This was the start of a systematic drive to ‘harmonise’ hate speech

laws beyond Holocaust denial across the EU. Or, to put it another way, this was a top-down diktat which aimed to counter any member states who might diverge in their attitudes towards dealing with racism and xenophobia.

Hate speech laws became a blunt instrument with which to beat member

states and their electorates into shape. As we will see, it could also be ignored to enforce compliance when it suited them as well.

Initially, the EU took a softly-softly approach to enforcing EU hate speech laws in member states. After decades of slow progress, they abandoned this and began legal proceedings against transgressors. In October 2020, the European Commission announced the initiation of infringement proceedings against Estonia due to its failure to transpose the Framework Decision entirely using criminal law. In its letter, the Commission pointed out that Estonia had not correctly criminalised hate speech by omitting the criminalisation of public incitement to violence or hatred when directed at groups and had not legislated for the provision of adequate penalties.⁶² Finland has found itself in the dock for failing to clamp down on hate speech and for not specifically targeting Holocaust denial in 2021. Infringement notifications on hate speech law implementations have been served against Luxembourg, Germany, and Hungary.⁶³

The recent attempt to make hate speech an EU federal crime – a move deliberately initiated to circumvent member state governments – shows how their patience with member states' slowness or reluctance to 'harmonise hate speech laws' across the Union had finally run out. In a statement justifying the attempt to include hate speech and hate crime Under Article 83 of the 'Treaty of the Functioning of the EU', they argue the following:

*Member States' criminal laws deal with hate speech and hate crime in different ways and that minimum harmonised rules at EU level exist only when such crimes are committed against a group or individual based on their race, skin colour, religion, descent or national or ethnic origin, which makes it difficult to implement a successful common strategy to effectively combat hatred.*⁶⁴

By a 'successful common strategy to effectively combat hatred', the Commission means it desires the ability to impose more stringent hate speech rules without the inconvenience of having to do this through member state governments and their electorates. The signalling of intent is a critical point to note because, for now, it is unlikely to be agreed upon by the Council of Ministers, which requires a unanimous vote. The direction of travel is openly displayed: Brussels will no doubt sidestep the member states in pursuing its campaign against free speech in Europe. After all, they have done so on numerous occasions.

The EU has made free speech a historical curiosity in slightly over one generation. The EU's Rule of Law through the Framework now upholds freedom from speech as a principal tenet. But what sped this process up enormously was the advent of the Internet and the emergence of social media.

The EU has made free speech a historical curiosity in slightly over one generation

The political seismic shock of Brexit and the election of Donald Trump brought home to Brussels how the battle over controlling the narrative was now a hot

war that they were determined to win. The true purpose and function of hate speech laws now came into sharp relief.

4.4 The Internet and Brexit

One of the most interesting things is how early the EU leadership recognised the dangers the Internet represented to their ability to control the EU narrative. The first clear warning of what the Internet would mean for the future was expressed by the Council of Europe when deliberating the implications of Cybercrime and the criminalisation of racism and xenophobia 'through computer systems' in 2003. They noted that

*the emergence of international communication networks like the Internet provides certain persons with modern and powerful means to support racism and xenophobia and enables them to disseminate easily and widely expressions containing such ideas.*⁶⁵

The fact that this observation was made in a discussion about cybercrime demonstrates how pervasive a bureaucratic sensitivity existed at the time to the speech implications of what the Internet would mean for the future. Brexit (and the election of Donald Trump in the USA), both of which were attributed to the spread of fake news and disinformation online, brought

A bureaucratic sensitivity existed at the time to the speech implications of what the Internet would mean for the future

this home to them with a vengeance. Their response was the Code of Conduct Countering Illegal Hate Speech Online enacted in 2016,⁶⁶ which came into operation without passing through the European Parliament. This was to have

far-reaching consequences, especially in preparing the ground for the Digital Services Act of 2024, a law passed to bring an existing practice within the EU's Rule of Law framework. That chronology is important to keep in mind.

The Code of Conduct set the scene for the birth of the largest censorship Trojan Horse on the planet.

5 The EU’s “operating system” for controlling speech

This chapter examines the EU’s system for controlling speech in detail. What is most remarkable about this system is the way it integrates EU institutions, NGOs and Big Tech into one unholy alliance of unaccountable regulators of speech. The degree of integration between these organisations is remarkable, and represents a profound threat to free speech in Europe.

5.1 The EU’s code of misconduct – the invisible hand of outsourced technocratic censorship

It is important to remember that the legal instrument underlying the Code of Conduct Countering Illegal Hate Speech Online remained the Framework Decision of 2008 discussed above. However, the Code of Conduct was not a piece of legislation. It was a voluntary code struck behind closed doors between the European Commission, the Big Tech companies, and a coterie of unelected NGOs self-designated as experts in hate speech and, thus, as ‘trusted flaggers’.

The Code referred to the framework for removing (without due process) what the Commission felt to be the illegal spread of hate speech online. However, it was also an experiment running public campaigns to increase ‘tolerance and pluralism’ online. The scope of its activities can be gleaned by

reading the Commission's 2019 'Assessment of the Code of Conduct on Hate Speech Online – State of Play'.⁶⁷

Without recording what content they refer to, it is impossible to know precisely what they designated as hate speech. However, the Code of Conduct claims to have resulted in the removal of 72 per cent of 'hate speech content' in 2019 (up from 28 per cent in 2016); 89 per cent of notices were reviewed within 24 hours in 2019 (up from 40 per cent in 2016).

The assessment also reveals how the EU Commission, along with Big Tech and a veritable army of 'trusted flagger' NGOs, grew in this period: since signing the Code of Conduct, for example, Twitter (now X) had enrolled 73 new trusted flagger organisations; YouTube had a four-times more extensive network of trusted flaggers compared to 2016, moving from 10 to 46 NGOs, and Facebook had increased its network of 9 partners in 2016 to 51 in 2019.

However, the work of the signees of the Code of Conduct went way beyond removing 'hate speech' online. For example, between 2017 and 2019,

The Code referred to the framework for removing, without due process, what the Commission felt to be the illegal spread of hate speech online

three workshops took place at YouTube, Twitter and Facebook headquarters to facilitate their 'tolerance and pluralism' goals. Most sinister of all was the

launching of an EU-wide online campaign by 40 NGOs in 24 languages during the 2019 European elections focused on promoting healthy and tolerant conversations online under the hashtag #'WeDeserveBetter'. This campaign reached over 6 million users on Facebook and Twitter. A 2018 pilot exercise to test such a campaign reached over 2 million users in several member states. Microsoft started a partnership with expert think tanks like the Institute for

Strategic Dialogue 'on counter speech to assist NGOs to surface and serve impactful counter narrative content via advertisements on Bing'.

The report also reveals a development we will return to in chapter 6, which is how Big Tech began to use automation and machine learning with help from human 'fact checkers' to enforce hate speech laws and the removal of posts deemed to be 'disinformation.' Meta, for example, cavalierly boasted that the interaction of human decision-making to train their technology,

All the players from the EU Commission onwards are all unelected and thus unaccountable to the electorate for their actions

which then 'becomes more accurate', had made their ability to police the online space more efficient.⁶⁸ In its Community Standards Enforcement Report (for the third quarter of 2021), the company said its proactive removal rate for hate speech

was 96.5 per cent. During the reporting period, 22.3 million pieces of hate speech were removed. YouTube's enforcement report (Q3 of 2021) records that between July 2021-Sept 2021, 5,901,241 videos were removed through AI moderation, 233,349 from user moderation, 85,791 from trusted flaggers, 9,471 from NGOs and 30 from governmental agencies.⁶⁹

The expansion of this censorship and disinformation machine is notable for two critical reasons: what unites all the players is the fact that from the EU Commission onwards, they are all unelected and thus unaccountable to the electorate for their actions. Second, this process circumvents member states and their electorates.

These dimensions of the hate speech narrative are hardly touched upon. However, they are a core component because in bypassing Parliament and member state scrutiny the EU effectively outsources control over what speech is allowed in Europe. This extra-legal alliance of unelected and unaccountable entities and organisations was not an accidental occurrence brought on by

the panic caused by Brexit. It was precisely how the EU bureaucracy operates in pursuit of its goals.

The arms-length relationship of the EU Commission to this process is its hallmark. This enables the EU Commission to maintain dispassionate innocence. When these practices were formalised through the passing of

The arms-length relationship of the EU Commission to this process is its hallmark. This enables the EU Commission to maintain dispassionate innocence

the DSA, Brussels, in the face of accusations of censorship, could merely defer to upholding the rule of law. But it also accomplishes something equally important: the extra-legal means of extending the extra-territorial scope and reach of the law without any parlia-

mentary oversight or accountability. Enforcing hate speech laws through extra-legal mechanisms outside of any democratic forms of accountability demonstrates that these laws are a means to another end rather than what they purport to be.

The Code of Conduct experience demonstrates that the EU bureaucracy was increasingly determined to impose its priorities on Europe without any pretence of upholding democratic accountability. The quest to determine the narrative trumped due process or even the pretence of a process. The means of operating through unelected and thus unaccountable corporations and NGOs became the EU's operating system. This is the Brussels 'Ministry of Truth' Josep Borrell is so keen to pretend does not exist.

5.2 The EU 'Ministry of Truth' – the unholy alliance of unaccountables and the privatisation of state censorship

The relationship between the EU Commission and unelected bodies like NGOs, corporations and the Courts is the backbone of how the EU

has institutionalised freedom from speech as a foundational value in Europe. The anti-democratic impulse behind this is what unites these players.

The role of the unelected NGOs is critical. They are fully integrated into the machinery of government in Brussels, mainly through the Commission's Register of Commission Expert Groups.⁷⁰ This consists of hundreds of

The relationship between the EU Commission and unelected bodies is the backbone of how the EU has institutionalised freedom from speech as a foundational value in Europe

unelected and unaccountable organisations and experts. According to the website,

the Expert Groups, which exists to 'ensure transparency', helps the Commission prepare 'legislative proposals and policy initiatives' and, among other things, with the 'implementation of EU legislation, programmes and policies, including coordination and cooperation with EU Member States and stakeholders'. In short, they play a significant role in how the EU creates, monitors, and executes its legislative programmes.

For our focus, the High-Level Group on Combating Hate Speech and Hate Crime (E03425)⁷¹ is an important group to examine. It is tasked with carrying out the above mandate concerning hate speech laws. Like all expert groups, their meetings are not open to the public, and there is no web streaming. Edited minutes of meetings are available to the public through the EU's cavernous website, which essentially hides their deliberations in public.

The structure is instructive: there are 27 member state authorities and five NGOs: Amnesty International European Institutions Office (AI EIO), the European Network Against Racism (ENAR), the Open Society European Policy Institute (OSEP), the Platform of European Social NGOs (Social Platform) and The European Region of the International Lesbian, Gay, Bisexual, Trans and Intersex Association (ILGA-Europe). In addition,

there are three other public entities: the European Commission against Racism and Intolerance (ECRI), the European Union Agency for Fundamental Rights (EU Agency) and the Organisation for Security and Co-operation Office for Democratic Institutions and Human Rights (ODHIR).

Not surprisingly, every deliberation of this committee ended with the conclusion that more restrictions on speech are required, not less. Every report ends with the need for more research or the 'clear need for more hate crime and hate speech training throughout the EU.'⁷²

This is not accidental. The self-serving tendency to expand the remit of these bodies institutionalises an expansionary dynamic which coincides with the EU's needs.

There are two aspects of the institutionalisation of expansion to be noted. The first is that the unaccountability of these entities means there are

The self-serving tendency to expand the remit of these bodies institutionalises an expansionary dynamic which coincides with the EU's needs

no external constraints on their activities. They attain a free-floating dynamic driven purely by the self-serving

interests of those involved. Their accountability to the EU Commission, which is, after all, keen to restrict free speech, is a blank cheque.

The second is that this extra-legal process inevitably results in the widening of hate speech laws while reducing the thresholds of what constitutes criminality. There are no constraints on the remit of its experts who are comfortable straying beyond their areas of expertise as they seek to ensure the relevance of their work.

The example of an organisation like the European Commission Against Racism and Intolerance (ECRI),⁷³ one of the expert Groups mentioned above, is paradigmatic.

Its role was pivotal after integrating new member states from central and Eastern Europe in the 1990s. Monitoring racism became part of the surveillance of these new countries, who, for different historical reasons, still retained attachments to their national sovereignty. ECRI's role was to weaponise hate speech laws as a counterpoint to any resistance to the EU's 'New Europe'.

As early as 2002, ECRI had already gone beyond its original remit on combatting racism. It now argued that criminal law should be used to penalise 'public insults' on the grounds of religion,⁷⁴ and that criminal law should penalize this and 'defamation' when committed intentionally. In 2013, ECRI expanded its mandate to cover 'hate speech ... against LGBT people.'⁷⁵ By 2015, ECRI had extended the scope of hate speech crimes to include 'stereotyping' and 'insults'. Their General Policy Recommendation No.15 on Combatting Hate Speech now expanded hate speech to encompass the following:

*The advocacy, promotion or incitement, in any form, of the denigration, hatred or vilification of a person or group of persons, as well as any harassment, insult, negative stereotyping, stigmatization or threat in respect of such a person or group of persons and the justification of all the preceding types of expression, on the grounds of "race", colour, descent, national or ethnic origin, age, disability, language, religion or belief, sex, gender, gender identity, sexual orientation and other personal characteristics or status.*⁷⁶

However, the recommendations ECRI drew from its massively expanded definition are more disturbing. They called for 'speedy reactions by public

figures to hate speech; promotion of self-regulation of media; raising awareness of the dangerous consequences of hate speech; withdrawing financial and other support from political parties that actively use hate speech; and criminalising its most extreme manifestations while respecting freedom of expression'.⁷⁷ Without a shadow of self-awareness, ECRI seriously suggests that clamping down on hate speech in the media and political debate is now a defence of freedom of expression.

ECRI is an NGO. This might be regarded as the musing of a self-appointed organisation that knows no limits. But this would be naïve in

Even though ECRI has no public or officially stated mandate for its pronouncements, it impacts the implementation of these laws

the extreme. Even though ECRI has no public or officially stated mandate for its pronouncements, and the monitoring of hate speech laws in member states is not sanctioned

in law and carries no binding commitment for member states, it impacts the implementation of these laws. The league tables of 'international standards', which ECRI arbitrarily creates, are wielded like a blunt instrument to measure member states' implementation of hate speech laws and pressurises them to generate more convictions and extend their criminal law. The result is the extra-legal and extra-territorial expansion of hate speech laws.⁷⁸

ECRI is one example. Another notable example, again part of the NGO umbrella organisations in the High-Level expert group noted above, is the International Lesbian, Gay, Bisexual, Trans and Intersex Association (ILGA). It produces the ILGA-Europe Rainbow Index,⁷⁹ which ranks countries according to their legal and policy practices concerning LGBTQI+ people. Again, this has no mandatory powers; it is simply a tool to pressure member states to enact hate speech laws more rigorously. Not surprisingly, Belgium scored high on the Rainbow Index in 2022, when the Belgian Justice Minister,

in response to an apparent increase in crimes based on sexual orientation, including discrimination, hate speech and hate crimes,⁸⁰ presented a plan to make the country a safer place for the LGBTQI+ community.⁸¹

The role of the European Court of Human Rights is even more important. Through its rulings and interpretation of hate speech laws, it not only sets a precedent but expands the scope of hate speech laws. A few disturbing cases demonstrate this.

5.3 The case of Sanchez v France (2021)

This case demonstrates how court rulings are expanding the scope of hate speech laws while significantly lowering the threshold of what constitutes criminality.

The case refers to a ruling that found a parliamentary candidate for Front National guilty of incitement to hatred and violence during an election through posts on Facebook. The court was unequivocal in its decision that posts on Facebook ‘clearly encouraged incitement to hatred and violence

Sanchez was found guilty as a proxy: it was not him who made the impugned comments but other Facebook users commenting on his post

against a person because of their belonging to a religion’. However, before examining their reasoning, it is important to realise that Sanchez was found guilty as a proxy. It was not him who made the impugned comments

but other Facebook users commenting on his post.

For the ECtHR, by allowing his Facebook wall to be public, Sanchez ‘assumed responsibility’ for the content of the comments posted. It agreed with the French court, which found that Sanchez left the comments up for six weeks before removing them and was thus guilty as the producer of an online

public communication site and, therefore, the principal offender – even though he had posted asking people to moderate their comments.

A previous finding in the case of *Delfi v Estonia* about content moderation of a news portal enabled the Court to reach this conclusion. The ECtHR simply transposed this ruling onto Sanchez because of his responsibility as a political candidate. To this end, the Court held that, although (in theory)

The curbing of freedom of expression in an election can be justified by an unelected and unaccountable European law court as being in the interests of defending democracy

political parties enjoy wide freedom of expression in an electoral context, this did not extend to racist or xenophobic discourse and that politicians had a

particular responsibility in combatting hate speech. So, although the speech was not his own, this did not appear to affect the court's decision to find no violation of Article 10, even though a criminal penalty had been imposed on the applicant. When this was later appealed to the Grand Chamber in May 2023, this Court also found no violation of Article (10)⁸² had occurred.

Article 10 says:

Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers.

In agreeing with the domestic court's decision that 'in the specific context of a forthcoming election', the comments 'could be classified as hate speech, when interpreted and analysed in terms of their immediate impact and were therefore unlawful'. The interference with the applicant's freedom of expression not only pursued 'the legitimate aim of protecting the reputation or rights of others, but also that of preventing disorder or crime' even though

there was no direct threat of violence or disorder. Its conclusion declared that ‘the interference in question could thus be regarded as “necessary in a democratic society”.’ There had therefore been no violation of Article 10 of the Convention. In short, there is no freedom of expression.

Remarkably, the curbing of freedom of expression in an election that does not incite violence can be justified by an unelected and unaccountable European law court as being in the interests of defending democracy!

But what is more chilling is some of the other reasoning of the Court. Referring to previous precedents,⁸³ which determined what constitutes incitement to hatred, it added the following curious explanation:

(this) did not necessarily require the calling of a specific act of violence or another criminal act. Attacks on persons committed through insults, ridicule or defamation aimed at specific population groups or incitation to discrimination, as in this case, sufficed for the authorities to give priority to fighting hate speech when confronted by the irresponsible use of freedom of expression which undermined people’s dignity, or even their safety. (my emphasis)

The Court effectively equated speech as violence even when it is not being used to incite violence. The ‘irresponsible use of freedom of expression’ begs the question of what ‘responsible freedom of speech’ might be. Now insults and ridicule, especially when used to undermine ‘people’s dignity’, are akin to calling for a pogrom. Undermining people’s dignity or even their safety

(without direct threats of violence) is now enough to justify the use of criminal law, especially given that the applicant, who

The Court effectively equated speech as violence even when it is not being used to incite violence

was not the creator of the comments, should be criminally liable for the content of others.

This ruling sets an astonishing precedent: speech is now violence. If the subjective test is met that words have undermined dignity, it is not only speech that is being policed but thought too, because to prove this, the speaker's intention must be referred to. After all, what is the objective threshold for determining the impugning of dignity?

This ruling also suggests that the use of social media in an election turns all posters of content into publishers who are now held to the standards of media publishers. Social media platforms do not fit in this scenario. Not only is the Court extending the scope of hate speech laws, but they are also sanctioning the policing of language and thought, effectively extending and lowering the threshold of criminality to such an extent that free speech is almost illegal – in the name of defending democracy.

5.4 The case of *Beizaras and Levickas v Lithuania* (2020)

What is unique about this case is that it differs from most other cases of hate speech as it emanates from the victim of the speech rather than the

What is unique about this case is that it emanates from the victim of the speech rather than the speaker

speaker. It is also a clear example of how the EU hierarchy imposes its values on a recalcitrant member state. This case involved hate speech on the grounds of sex and sexual

orientation in Lithuania, which used the elevation of LGBTQ victimhood to make its ruling.

Notably, the case came before the Court via a joint action of the defendants and ILGA,⁸⁴ which we previously noted, is one of the EU Commission's Expert Groups in the sphere of hate speech and 'fundamental rights'. Not surprisingly, because the ILGA argue that 'LGBT people are

perceived as one of the most vulnerable social groups in Lithuania',⁸⁵ they saw this as an important case to pursue.

The case centred around a post on Facebook by one of the applicants which was a photograph of him kissing his male partner, the second applicant. This resulted in hundreds of hateful and discriminatory comments under the post. The applicants held that the Lithuanian authorities' refusal to launch a pre-trial investigation into the comments left under the post amounted to a violation of Article 8 of the European Convention of Human Rights, which states that 'Everyone has the right to respect for his private and family life, his home, and his correspondence'. Furthermore, the applicants argued that this also violated Article 14, which is based on the core principle that all of us, no matter who we are, enjoy the same human rights and should have equal access. The Court agreed with them.

The press release explaining the ECtHR's decision to uphold the applicant's case states the following:

*The Court found in particular that the applicants' sexual orientation had played a role in the way they had been treated by the authorities, which had quite clearly expressed disapproval of them so publicly demonstrating their homosexuality when refusing to launch a pre-trial investigation. Such a discriminatory attitude had meant that the applicants had not been protected, as was their right under the criminal law, from undisguised calls for an attack on their physical and mental integrity.*⁸⁶

It is curious how a defence can be made of a violation of privacy for public responses to a personal photograph posted on Facebook and shared with the world. However, such an anomaly is unimportant because the Court was intent on sending an unmistakable message about its official restriction on what it now broadly interprets as hate speech. The right to free speech,

especially when it is upheld by a national Court perceived to be anti-LGBTQ, is trumped by protecting a minority being 'victimsed'.

This effectively institutionalises a new hierarchy of freedoms based on victim-based law: the right to a private life now trumps the duty to uphold free speech in public. Ironically, this nullifies the right to free speech and effectively creates a two-tiered system of legality where those deemed to

The right to free speech is trumped by protecting a minority being 'victimsed'

breach hate speech, even when there is no immediate threat of violence, do not enjoy the 'same human rights' and certainly do not have 'equal access

to them' as the HRC contends. The right to privacy, especially for groups perceived to be vulnerable, now trumps the right to free speech.

This case is an example of a precedent that expands the applicability of hate speech laws where a victim's perception of hate (which, therefore, assumes they know the intention behind the use of words) is the basis for criminalisation. This is policing thought and language. The right to freedom from speech formally replaces the right to free speech.

5.5 Big Brother and Big Tech

However, one more dimension is equally important, and that's the role played by Big Tech, through which the EU Commission outsources censorship.

The outsourcing of the policing of speech in the 21st-century public square – online – has been systematised by passing the DSA into law.

We saw above what impact the Code of Conduct had on the removal of 'hate speech infringements'. But the replacement of the Code by a Law marks a massive step up in intent. The financial penalties and time pressures instituted through the DSA mean the direction of travel is a foregone conclusion. A quick examination of the DSA Data Transparency Database⁸⁷ –

the database where the social media platforms share their content moderation

This is privileged information, available only to regulators and selected researchers, not the public

actions – shows how vast the takedown of hate speech-related content has become. Unfortunately, this database, which we are told ensures ‘transparency’, doesn’t reveal the content of

‘hate speech transgressions’. This is privileged information available only to regulators and selected researchers, not the public. However, by examining how these Big Tech platforms conceptualise hate speech, it is not unreasonable to assume that the scope of what constitutes grounds for removal has been widened, which significantly lowers the threshold of what speech criminality now consists of.

Today, Facebook and Instagram, for example, conceptualize hate speech in the form of a ‘direct attack’ based on an extensive list of protected characteristics such as race, ethnicity, sexual orientation, caste, and serious disease. Such an attack is broadly defined as ‘violent or dehumanizing speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing and calls for exclusion or segregation’.⁸⁸ Other platforms have also adopted this remarkably lowered threshold of permissible speech.

Other platforms have also adopted this remarkably lowered threshold of permissible speech

YouTube now extends the prohibition to slurs and stereotypes; TikTok also polices slurs; and the former Twitter prohibits, amongst others, slurs, epithets, racist and sexist tropes or other content that degrade

people. It is impossible to see how such lowered thresholds of permissible speech will not impede legitimate debate. Simply questioning immigration policies, the environmental impacts on communities or the desirability of

a Federal Europe must inevitably fall foul of such low thresholds. But this is the point, not the flaw.

What becomes clear when one takes all the activities of the ‘unholy alliance of the unaccountables’ into account is how far away from the defence of free speech the EU has travelled. Their insulation from any blame for the systemisation of censorship in Europe is necessary to maintain a pretence they still uphold democracy. However, hidden from view and outsourced,

It is impossible to see how such lowered thresholds of permissible speech will not impede legitimate debate

all these entities bask in the shadow of the rule of law, where each player absolves itself from any responsibility for the authoritarian assault on free speech.

The ‘unholy alliance of unaccountables’ is how the EU Commission becomes untouchable. Parroting Jim Malone, the cop played by Sean Connery in the movie, ‘The Untouchables’, this is not the ‘Chicago Way’ but the ‘Brussels Way’.

The warning by Borrell, in the speech referenced in the introduction, that the EU must not slip into censorship, that it ‘cannot invent the “Ministry of Truth”’, which says ‘this is true, this is not true’, is a remarkable example of Orwellian double-speak.

From what we have described above, the evolution of hate speech laws reveals that a de facto EU ‘Ministry of Truth’ has not only come into being but is flourishing in Brussels.

The DSA is just one strand in its armoury. Together with the new Media Freedom Act (MFA),⁸⁹ the Child Sexual Abuse Regulation (CSAR),⁹⁰ and the Artificial Intelligence Act⁹¹ to legislated restrictions on political advertising,⁹² the EU has systematically built a legal citadel, a ‘censorship operating system’ through which to institutionalise freedom from speech in Europe.

5.6 The institutionalisation of freedom from speech

The evolution of hate speech laws in the EU can now be seen as a steady and systematic process which inexorably leads to the ending of the right to free speech as a foundation of European democracy.

The EU has always been a top-down phenomenon, not an organic development rooted in popular sovereignty. Since its inception as the European Coal and Steel Community in 1953, then the European Economic Community from 1956 to the European Union since 1993, the only consistent value endorsed by the EU elite has been 'anti-democracy – the creation of a system that separates power and control in Europe from any expression of the popular will'. It is the anti-democratic union of Europe's political elite.

This top-down imposition on the people of Europe poses a constant legitimacy problem for the EU elite because their project is defined more by

Driven by the inner needs of the elite, the attack on free speech has become a political and moral crusade

what they are against rather than what they are for. This is the source of a foundational crisis and a defensiveness that permeates the entire body politic of the EU's leaders and institutions. It manifests itself in the constant need to break Europe from its historical attachments to sovereign nation-based cultures and values to forge a new European identity and values

that reflect its cosmopolitan federalist outlook and aspirations. Controlling and expanding this narrative is the omnipotent existential dilemma facing the Eurocrats.

The inner needs of the elite mean they are constantly casting around for legitimising moments to gain authority for their rule. This explains why the attack on free speech has become a political and moral crusade. This crusade is not simply about cleansing Europe of hateful speech. Its fundamental purpose is to promote political and social change. The object

is not merely to police language but to control and constrain verbal communication. The dynamic is not just to police speech but increasingly to change how Europe's citizens think and regard meaning. It is a technocratic and authoritarian attempt to socially engineer outcomes that serve the EU elite's interests.

This crusade can never end. The accomplishment of this mission is impossible. Policing language to alter behaviour and social outcomes can

The disinformation narrative about malign foreign intervention in the forthcoming elections means the drive to police speech will only intensify

never be achieved. The reasons are simple: first, you cannot outlaw ideas and thoughts. Second, there is the pesky issue of the demos and their democratic rights. As long as there are people

in Europe who have inconvenient rights like the right to vote in elections – people the EU elite think are incapable of independent thought – this political and moral crusade is forced to continue, to expand and up the stakes.

Thus, during the discussion of extending the list of EU crimes to hate speech and hate crime, the Commission was urged by the Parliament to consider:

*an open-ended approach whereby the list of grounds of discrimination will not be limited to a closed list in order to effectively combat hate speech and hate crimes motivated by new and changing social dynamics.*⁹³

The open-ended changing social dynamics, now apparently, should include minors, 'including those belonging to vulnerable groups' to provide them with special protection from hate speech and hate crimes. The High-Level Expert Group on Hate Speech now wants to include the psychological damage 'intersectional hate crimes cause individuals and communities'.⁹⁴

The lumping together of children with vulnerable groups is a deliberate ploy to justify an 'open-ended' approach which would give Brussels *carte blanche* to apply hate speech and hate crime legislation to almost anything that moves in Europe.

The danger of this inner dynamic cannot be overstated. The disinformation narrative about malign foreign intervention in the forthcoming elections, which is consciously being wielded to smear the rising tide of support for populist 'far right' parties, means the drive to police speech will only intensify. A crucial part of this will be the weaponisation of hate speech detection by automated AI systems, which the EU are embracing with the enthusiasm of new converts. This is the focus of the next Section.

To suggest that freedom from speech has become the hallmark of European democracy is not a rhetorical flourish. While freedom of expression is often paid lip service to, the trajectory over the past quarter century has seen the expanding erosion of any attachment to upholding free speech. The systematic and consistent character of the open hostility to free speech reveals that there is a dynamic external to the issue of expression that is driving this behaviour. The willingness to use and abuse the Rule of Law to suit their purposes suggests that a systemic dynamic is operating, resulting in anti-democratic authoritarian behaviour. That dynamic is the crisis of authority at the heart of the EU project, which relentlessly drives the EU's determination to police the speech and thoughts of 448 million Europeans. Explaining how this dynamic operates is the focus of the next chapter.

6 AI and the automation of language policing

We saw from the previous chapter how linguistic policing to socially engineer outcomes that reinforce the EU's cosmopolitan federalism remains one of the critical drivers of the crusade to criminalise hate speech. One of the most disturbing developments which threatens to weaponise this process way beyond current capabilities is the development of Artificial Intelligence, particularly generative AI which is being used to automate the detection of hate speech online.

The celebration of passing the EU's Artificial Intelligence Act (AIA) in Parliament was a grotesque bout of self-congratulations. MEPs, and no doubt the Commission, were quick to herald the world's first regulation of AI. The press release⁹⁵ triumphantly announced that the AIA will protect 'fundamental rights, democracy, the rule of law and environmental sustainability from high-risk AI'. It will 'ban certain AI applications that threaten citizens' rights ... and AI that manipulates human behaviour or exploits people's vulnerabilities will also be forbidden'. During the plenary debate, the Internal Market Committee co-rapporteur Brando Benifei (S&D, Italy) stated that thanks to Parliament, 'unacceptable AI practices will be banned in Europe, and the rights of workers and citizens will be protected'. The Civil Liberties Committee co-rapporteur Dragos Tudorache (Renew, Romania) added that the EU had delivered because it 'linked the concept of artificial intelligence to the fundamental values that form the basis of our societies ... The AI Act is a starting point for a new model of governance built around technology. We must now focus on putting this law into practice'.

This all sounds wonderfully righteous. Except for the fact that AI has already been deployed in Europe as part of an older technocratic governance model – the automation of hate speech detection by social media platforms.

Unsurprisingly, the EU Commission, whose chief motivation for regulating AI is their risk-averse reverence for the precautionary principle,

This is genuinely the application of ‘high-risk AI’, which mortally threatens citizen’s rights and consciously aims to manipulate human behaviour

have no problem using this technology when it comes to policing language. If truth be told, this is genuinely the application of ‘high-risk AI’, which mortally threatens fundamental citizen’s rights and

democracy and consciously aims to manipulate human behaviour. But there is no precautionary principle operating here. If anything, they desire more AI-driven algorithmic automation because this serves their interests.

6.1 A true disinformation narrative

There is no doubt that generative AI will impact the forthcoming elections. However, suggesting that this is the gravest threat to European democracy is an Orwellian disinformation narrative that masks a real danger.

It is essential to begin this examination by taking apart an assumption repeated as if it is a self-evident truth, namely, that widespread disinformation and hate speech online negatively impact outcomes. This is an assertion, not a scientific fact.

A report by Harvard Kennedy School academics, ‘Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown’,⁹⁶ counters this narrative convincingly. The authors suggest that the consumption of misinformation is mainly limited by demand and not by supply. Numerous studies show that despite the quantity and accessibility

of misinformation, the average internet user consumes very little of it.^{97 98}

Where disinformation is consumed, it is usually heavily concentrated in

That widespread disinformation and hate speech online negatively impact outcomes is an assertion, not a scientific fact

a small portion of active and vocal users.⁹⁹ Interestingly, the researchers view partisanship and identity as key determinants of misinformation belief and sharing,

while the lack of access to reliable information only plays a negligible role.¹⁰⁰

Most people consume content from mainstream sources, typically the same handful of popular media outlets. Contrary to every assertion about the spread of hate speech and disinformation online, the report suggests that, on its own, it has no causal effect on the world. Most telling, they conclude, citing other well-researched reports, that such panics ‘might be based on the mistaken assumption that people are gullible, driven in part by the third-person effect’.^{101 102}

If there is any section of society that suffers from the ‘third-person effect’, it is the EU elite who believe that everyone, apart from themselves, are gullible and lacks the independent adult capacity to tell truth from lies. Never mind the fact that the dangers of hate speech and disinformation online are exaggerated, what this serves to justify is the green light for Brussels to adopt this same technology to fulfil its crusade against free speech. It is notable how, in all these discussions, the dangers of algorithmic manipulations, which are held responsible for the spread of fake news or hate speech, become righteous tools if recruited to support the crusade against hate speech.

The Civil Liberties for Europe and European Partnership for Democracy report ‘Identifying, analysing, assessing and mitigating potential negative effects on civic discourse and electoral processes: a minimum menu of risks

very large online platforms should take heed of’, January 2024,¹⁰³ is paradigmatic of this technocratic narrative which claims to ‘ensure robust protection of civic discourse and electoral processes under the DSA’.

For its authors, ‘illiberal forces advocating a purely majoritarian view of democracy have ascended to power and solidified their positions’ in Europe. This shift has, apparently, ‘led to a weakening of the rule of law, a foundational pillar of democratic governance’. So, anyone advocating a ‘purely majoritarian

Contrary to every assertion about the spread of hate speech and disinformation online, the report suggests that on its own it has no causal effect on the world

view of democracy’ – in short, anyone who believes in democracy where majorities determine outcomes – is now a threat to democracy. To counter such threats, the report focuses on ‘inclusive, pluralistic, and accessible ... algorithmic configurations in VLOPs (Very Large Online Platforms)

and VLOSEs (Very Large Online Platforms and Search Engines) that play a pivotal role in shaping the information landscape’. Their insights into the inherent dangers are drawn from an influential book, *Technology and Democracy: Understanding the Influence of online technologies on political behaviour and Decision-making*, published by the European Commission JRC Science for Policy Report 2020.

Not surprisingly, this book – a behavioural psychology perspective – argues that ‘social media changes people’s political behaviour (sic)’ and that ‘there is scientific evidence that social media changes people’s political behaviour offline’. And true to form, that ‘this includes the incitement of dangerous behaviours such as hate crimes.’ Contrary to reports that debunk the claim that algorithmic manipulations by social media platforms have little effect, they have ‘scientific findings’, which reveal that ‘algorithms that promote attractive, engaging content and people’s strong predisposition

to orient towards negative news, as most “fake news” tends to evoke negative emotions such as fear, anger and outrage.’ The most critical technocratic insight is the following:

*Scientific findings suggest that there is an ideological asymmetry in the prevalence of echo chambers, with people on the populist right being more likely to consume and share untrustworthy information.*¹⁰⁴

What alternative does Brussels have other than to hunt out these echo chambers, shut them down and stop the spread of ‘untrustworthy information’ that, God forbid, might seep into broader society and result

These champions of civic discourse who remain so hostile to technologies like social media, become fervent technocrats when it comes to manipulating these to further their ends

in ‘majoritarian democratic’ outcomes?

What is needed are mitigating strategies – authoritarian nudging strategies – for ‘creating a more inclusive, diverse, and accessible digital public sphere where varied voices can contribute to a healthier civic discourse’. And a ‘civic discourse’ they are so keen to encourage can only remain

uncivil if it encompasses ‘behaviours from impoliteness to hateful or even hate speech’. Never mind the paternalism and infantilisation with which this drips. The message is clear: police the online sphere for the wrong language or else. The Council of Europe’s report, ‘Assessment of the Code of Conduct on Hate Speech Online State of Play’ report 2019, mentions concepts like ‘exclusionary nationalism’ should be part of the hit list.¹⁰⁵

Remarkably, these champions of civic discourse who remain so hostile to technologies like social media, which they think determine outcomes, become fervent technocrats when it comes to manipulating these to further their ends. This is truly dangerous.

6.2 AI and automating authoritarianism

As we mentioned earlier, the discussion of the dangers of generative AI and algorithmic manipulations hides from view a more insidious application of this technology to automate hate speech and disinformation detection online. This has become a growth area. If you do a search on Semantic Scholar for ‘hate speech detection,’ you will see 37,400 results of papers with this in their title. In the past ten years, the number of papers has significantly increased: in 2000, there were 324 papers produced. By 2020, however, this rose to 2,806. Today, 118,000 articles on Google Scholar deal with AI and automated hate speech detection.

We have already alluded to the fact that social media platforms have been using these technologies to aid their detection of hate speech and disinformation. Facebook reported that in the first quarter of 2019, 65.4 per cent of the content removed was flagged by machines (with an increase from 51.5 per cent in the previous months). YouTube reports that, in 2017, 79 per cent of the videos removed for violating their policies were initially flagged by automatic flagging systems. In the second quarter of 2019, this was 87 per cent. Many of the removed videos are taken down before being viewed by users. By April 2019, through using technology, 38 per cent of abusive content that Twitter actions surfaced proactively for human review instead of relying on users’ reports. This marks a significant increase from the previous year, where machines flagged 20 per cent of potentially abusive content. It should be noted that all content surfaced by an automatic detection system is assessed by the reviewers before being actioned (human-in-the-loop).¹⁰⁶

There are two related dangers of this weaponisation of hate speech detection.

First, automating online hate speech detection is fraught with difficulties. As one study argues, ‘subtleties in language, differing definitions on what

constitutes hate speech, and data availability limitations for training and testing these systems¹⁰⁷ pose huge challenges. Understanding different contexts adds further complexity.¹⁰⁸ Another study reveals how AI has a ‘limited ability to parse the nuanced meaning of human communication, or to detect the intent or motivation of the speaker’.¹⁰⁹ Given these difficulties, it suggests that placing trust and authority in such systems, which can inadvertently label something as hate speech, is highly problematic.

But this is not a problem for EU technocrats. It is a bonus because such systems will widen the long arm of the Brussels censorship machine operating under the simple rule that it is better to be safe than sorry.

In short, the automation of hate speech detection using AI will necessarily lower the threshold of what constitutes hate speech. The inbuilt drive to expand the scope of hate speech detection means the EU’s crusade can now be fully automated and technologized, removing the human as the last bastion of reason in systemic censorship.

This is the core message of the European Agency for Fundamental Rights’ 2023 report ‘Online Content Moderation: Current Challenges in Detecting

Such systems will widen the long arm of the Brussels censorship machine, under the simple rule that it is better to be safe than sorry

Hate Speech’¹¹⁰ whose objective is to provide a ‘critical assessment of the limitations of online content moderation tools in detecting online hate against specific groups.’ They want to improve this, not dispense with it. They are

convinced that using AI will increase the efficiency of content moderation and ‘scale up tasks that would be impossible to undertake through human review alone’. They reveal that the automation of hate speech detection is already using an online lexicon including lists of words that indicate hate and malicious speech developed by the Weaponized Word database¹¹¹ and the

Dark Data Project – an unelected organisation supported by numerous unelected NGOs, UN bodies and entities like the Rand Corporation.

In the Foreword by Director Michael O’Flaherty, the assumptions which echo the EU’s concerns are spelt out clearly:

Tackling online hate speech is about protecting the rights of victims of hate speech ... It is imperative that human rights voices are centrally involved in the design and implementation of moderation measures ... Investing in a variety of measures to tackle hate speech is critical. Only then can we genuinely safeguard people’s rights online and create spaces for people to connect, learn, share thoughts and join discussions online.

Using automated AI systems to detect hate speech is not about protecting the right to free speech but protecting victims’ rights from speech. Moderation

Moderation systems are the weaponisation of language policing for the 21st Century

systems that are fully automated are now regarded as the modernised form of content moderation, based as they are, on lexicons drawn up by self-appointed,

unelected, and unaccountable organisations. This is the weaponisation of language policing for the 21st Century.

The second problem relates to the attack not just on free speech but on privacy, too. The concern that AI systems are not good enough and need to become more robust and accurate is leading to advocacy for enabling them to access more data and metadata of words and user behaviours online. For these systems to be more precise and effective, they need access to broader data sets besides an ontology, dictionary or text containing potentially hateful keywords. One report argues,

Additional information from social media can help further understand the characteristics of the posts and potentially lead to a better identifica-

*tion approach. Information such as demographics of the posting user, location, timestamp, or even social engagement on the platform can all give further understanding of the post in different granularity.*¹¹²

For the authors writing in 2019, the problem was that this kind of data was not publicly available to researchers. However, with the DSA, as we noted when examining the Transparency Database, this is no longer an issue. Despite the assurances of strict compliance for researchers gaining access to user data and the metadata, this information is now available and will no doubt be used to ‘improve’ studies on how online risks evolve and assess the effectiveness of platforms’ policies and measures against harmful content.

The report’s conclusion indicates there is only one direction of travel: *The need for automatic hate speech detection systems becomes more apparent ... Given all the challenges that remain, there is a need for more research on this problem, including both technical and practical matters.*

With the DSA, this only means the greater granulated identification of hate speech and those committing it. Indeed, this precise point was stressed during a panel discussion in Germany on the rise of the ‘far-right’.¹¹³

If you think this is a discreet debate among AI geeks, think again. A 2022 Council of European Committee of Ministers (CECM) recommendation on combating hate speech, for example, argues that:

*state authorities, national human rights institutions, equality bodies, civil society organisations, the media, internet intermediaries and other relevant stakeholders should not only cooperate on specific initiatives, but also share data and best practice, and, via coordinated mid-term action plans, work more thoroughly on prevention.*¹¹⁴

By ‘work more thoroughly on prevention’ the CECM is talking about making censorship more efficient. Indeed, one of the ‘stakeholders’ CECM cites in

this document is the EU-funded European Online Hate Lab (EOHL)¹¹⁵, whose goal is to improve detection and the EU's capacity to act. One of its primary objectives is to 'increase the knowledge of ecosystems of hatred online and the capacity to respond.' Rooting out 'ecosystems of hatred', which would require data on users and their networks, means weaponising the war on hate speech by destroying privacy.

The terminology of 'ecosystems' is not incidental. Automating hate speech detection requires information on social context, not just taxonomies

'Containment policies' which 'emphasize behavioural interventions' must sound like music to the ears of EU technocrats intent on socially-engineering outcomes

of words. As one study argues to justify its approach, they advocate the formalising of 'the

idea that an individual's hateful content is influenced by their social circle'. They 'propose a framework that combines text content with social context to detect hate speech ... Our results suggest that considering social context is a promising direction for improving hate speech detection on Twitter'.¹¹⁶ Once social context becomes a legitimate target for inclusion, it is a matter of time before user behaviour is in the crosshairs. Another study advocates this unsubtle approach, arguing that spreading hate speech and disinformation by malign intent needs 'containment policies' which 'emphasize behavioural interventions'.¹¹⁷ This must sound like music to the ears of EU technocrats intent on socially-engineering outcomes to provide social stability in Europe.

The shift this encompasses is significant to note, too. In the mid-1990s, 'Internet intermediaries' (today's Social Media Platforms) were governed by the 'immunity doctrine',¹¹⁸ which regarded them as technically neutral conduits, agnostic about users' content. But because of the DSA, the handling of hate speech has transformed yesterday's intermediaries into

the ‘new governors of online speech’.¹¹⁹ The legally sanctioned obligations to process complaints in hours or days¹²⁰ at the risk of fines ensures that the drive towards the automation of flagging and removing hate speech is overwhelming – for both the EU and Big Tech.

6.3 Brussels on the warpath

The automation of hate speech detection through the application of AI is placing the policing of speech in the forthcoming elections on a war footing. The recent Munich Security Conference (MSC) announcement of a ‘Tech Accord to Combat Deceptive Use of AI’ in the 2024 election¹²¹ is a case in point. Twenty leading technology companies, including Adobe, Amazon, Google, IBM, Meta, Microsoft, OpenAI, TikTok, and X, have pledged to

The automation of hate speech detection through AI is placing the policing of speech in the forthcoming elections on a war footing

work together to detect and counter harmful AI content. For MSC Chairman Ambassador Dr Christoph Heusgen, this is crucial in ‘advancing election integrity, increasing societal resilience, and creating trustworthy tech practices’. Reining in

‘threats emanating from AI while employing it for democratic good at the same time’ – automating hate speech detection – is now apparently the way to ensure democracy is defended.

Ambassador Christoph Heusgen is another member of the ‘unholy alliance of unaccountables’. He has never been elected to the positions of power he now occupies. His career is dazzling. He has been a Permanent Representative of Germany to the United Nations between 2017 and 2021, Angela Merkel’s Foreign Policy and Security Adviser, and the Director of the Policy Unit for High Representative Javier Solana in the General Secretariat of the Council of the European Union from 1999 to 2005.

However, not one voter in Germany or Europe has ever endorsed his exemplary credentials.

Part of the Brussels censorship war machine is its army of unelected fact-checking flagging organisations across Europe. Indeed, heading this up

Part of the Brussels censorship war machine is its army of unelected fact-checking flagging organisations across Europe

for the Commission is the small Irish Media regulator, Coimisiún na Meán,¹²² who, under its unelected Executive Chairperson, Jeremy Godfrey, will oversee

handling most of the regulatory DSA caseload for the entire bloc. As the EU's truly vassal colony in Europe, Ireland, through Godfrey, has been granted unprecedented powers to censor online content for 450 million EU citizens. His 75-person office will effectively become the chief EU moderator for platforms like Facebook, X (formerly Twitter), and TikTok, deciding which content to remove.¹²³

Big Tech is also mobilising itself in support of Brussels. Marco Pancini, the head of Meta's EU Affairs, announced that Meta has established a so-called Elections Operations Center to facilitate its monitoring efforts.

This will gather experts from its intelligence, data science, engineering, and

Big Tech is also mobilising itself in support of Brussels

legal teams to 'identify potential threats and put specific mitigations in place across our apps and technologies in real-time'.¹²⁴

It sounds very efficient but means high-speed 'real-time' censorship.

An authoritarian technocratic managerialism, which will mostly remain hidden from the public, is in operation. If you want an insight into how this will operate, examine the musings of Benoît Loutrel, a former Google Executive who has been appointed to head up France's audiovisual and

digital media regulator, Autorité de régulation de la communication audiovisuelle et numérique (ARCOM). ARCOM is part of the same DSA monitoring network headed by Ireland’s Coimisiún na Meán.

In an interview with *Politico*’s Chief Technology correspondent, Mark Smith, he makes an important point about the DSA, which few people have noticed. ‘We are bringing something new to what we have done in the past’, he states. ‘It’s a new model. For the first time, we are not the French regulator. We are the French player of the EU team of regulation.’¹²⁵

This unelected official is the EU’s, not France’s enforcer, even though France has not yet passed the necessary national legislation to make enforcement legal for its European election. But having no legal basis yet has not stopped ARCOM from holding a meeting with social media companies, civil society groups, government ministries and the organizing committee for the upcoming Olympic Games in Paris to develop a ‘playbook to safeguard the global event from hateful and illegal content’ reports Smith in his newsletter.

It is Loutrel’s cavalier technocratic mentality that is truly disturbing. He sounds very reasonable when he states that he is not interested in issuing blockbuster multibillion-dollar fines under the Digital Services Act. But why not? Because, as he argues, this would be a sign of failure. Instead, he is more

‘Success of the DSA is not that we fine people ... success is that you get compliance’

interested in regular behind-the-scenes communications with companies to ensure they meet their obligations. ‘Success of the DSA is not that we fine people,’ he tells Smith, ‘Success is that you get compliance’. In other words, issuing fines would mean that social media platforms are not acting with due haste and are not removing material deemed ‘hateful’ or ‘disinformation’, according to Brussels.

This might sound logical, but it is genuinely sinister. In his view, not issuing fines – the measure of the success of enforcement – means the policing of online speech and disinformation remains hidden from the public.

It is almost medieval in its logic: the only time we hear of a ‘crime’ is when the officials announce a public hanging

It is a discrete and increasingly automated process which may or may not involve dialogue between unelected officials like Loutrel and Big Tech. Fines, on the other hand, which he wants to avoid, would be one of the only ways in which the public

would be made aware that censorship is taking place. It is almost medieval in its bureaucratic and technocratic logic: the only time we hear of a ‘crime’ is when the officials announce a public hanging.

As noted in the previous chapter, the beauty of all this is that it provides Brussels with a censors get-out-of-jail card. After all, it is not Brussels that has taken down the millions of alleged hate speech posts, but Big Tech. By automating these systems, Big Tech can claim it is merely following orders and the rule of law, a necessary price to pay for doing business in Europe.

However, the net effect is that the threshold of what constitutes hate speech has been substantially lowered while accountability for Brussels’s censorship will be automated out of existence.

However, the most pernicious thing about the long-term weaponisation of language policing through AI is that it automates semantic engineering. The granularity required to make these systems robust means the words used and how they relate to each other, especially for Large Language Models, need to ‘contextualise’ speech. This means the hate speech crusade now consists of engineering not just words and speech but the thoughts they express. In the hands of the EU’s ‘Ministry of Truth’, this is an Orwellian nightmare in the making.

Automated semantic engineering will develop a taxonomy of hate, a toxic expanding list of words that must be quarantined and isolated. This sends out a strong message to the real target of the EU elite's ire, the public. Insulating the public from toxicity is not a benign exercise of ensuring social harmony. It is a hateful narrative that transforms European citizens from rights-bearing agents into the object of an epidemiological intervention to protect them from the speech the elite believe they have no capacity to understand.

Finally, it institutionalises the EU elite's right to determine what speech is acceptable or not. It legitimises the criminalisation of anyone, particularly

Running an election not based on free speech but increasingly automated freedom *from* speech is how the EU elite are trying to address their crisis of authority

the 'far right', who might have contrary views about the importance of national sovereignty or who oppose woke identity on perfectly reasonable grounds. The

underlying dynamic dictates this as an inevitable outcome. Running an election not based on free speech but increasingly automated freedom from speech is how the EU elite are trying to address their crisis of authority. The stakes could not be higher, which is why the battle of the narratives is so critical. This is the focus of the following and concluding chapter of the report.

7 The forthcoming elections and the battle of the narratives

In the introduction, we referred to the recent speech by the EU High Representative of the Union for Foreign Affairs and Security Policy/ Vice-President of the Commission, Josep Borrell, in which he warned about the ‘battle of narratives’ in the forthcoming EU elections. He was right. This is a battle that democrats must win against Eurocrats such as Borrell.

The Brussels machine has already launched its campaign, as demonstrated in the panic about disinformation over the current farmers’ protests. Through *EUvsDisinfo* services¹²⁶ – the flagship project of the European External Action Service’s East StratCom Task Force, part of the EU’s diplomatic services under Borrell – they have attempted to delegitimise the farmers by linking them to Russian ‘disinformation’. The EU’s narrative is that Moscow’s ‘disinfo peddlers’ are trying to ‘exacerbate the perception of divisions between the people, and what the Kremlin erroneously calls the ‘Brussels elites’.¹²⁷ By referring to the ‘perception of divisions’ they suggest that the open conflict between Europe’s farmers and Brussels is a perception rather than a reality. The scary inverted commas around the phrase ‘Brussels elites’ further questions this reality. In short, the narrative operates on two levels: first, it links Kremlin disinformation to an actual popular political

movement threatening the EU machine, in order to discredit it. Second, it reinforces the alleged dangers of foreign intervention in the public mind.

The reported discovery by French authorities of a ‘vast Russian disinformation campaign in Europe’ has been similarly used to reinforce the couplet of Russian interference and political divisions in France.¹²⁸

This is disingenuous. The EU bureaucracy has no problem with foreign involvement in the internal affairs of the EU when it serves its crusade against hate speech. As Professor Alexander Peukert of Goethe University, Frankfurt am Main, points out, many of the trusted flaggers of illegal content under the DSA are US-based actors who do not comply with the DSA’s stipulation that these entities should be established in member states.¹²⁹

Take the case of one of the Commission’s High-Level Expert Groups, the European Region of the International Lesbian, Gay, Bisexual, Trans and Intersex Association (ILGA), which we discussed in chapter 4. It is curious that this registered Expert Group which received €464,301 from the Global Equality Fund (GEF)¹³⁰ in 2022 (and €1 million from the Commission)¹³¹ does not feature among the concerns in Brussels about foreign intervention in the internal affairs of the EU. The GEF is managed by the U.S. Department of State’s Bureau of Democracy, Human Rights and Labor. But, of course, because the GEF is ‘dedicated to advancing and defending the human rights and fundamental freedoms of lesbian, gay, bisexual, transgender, queer and intersex (LGBTQI+) persons around the world’, it’s perfectly acceptable and desirable that this foreign influence is at the heart of EU policymaking. Anything that advances the crusade against hate speech is good; anything that doesn’t or questions it is bad.

This is the Orwellian ‘Brussels Way’ in action. It is prevalent in every sphere of activity through which Brussels attempts to control and dictate the narrative.

7.1 The European Media Freedom Act (EMFA)

This act¹³² is an important example of Brussels’ Orwellian double-speak. As a law purporting to encourage media freedom and pluralism, the EMFA centralises Brussels’ control over all national media activities in the EU. It is a fig leaf for the EU’s crusade to undermine the sovereignty of European nations. The EU oligarchy has long had a problem with upstart or dissident member states, whom Brussels seeks to control by erroneously alleging that they pose a risk to European media freedom.¹³³

When the EMFA was first proposed in September 2022, it was against a backdrop of fearmongering about a supposed crisis of democracy in Central Europe. Vice-president of the European Commission for ‘values

Brussels is relaxed about media monopolies in other EU member states: according to the Media Pluralism Monitor (MPM) no European country is free from threats to media pluralism

and transparency’, Věra Jourová, explicitly mentioned Hungary in the EMFA: ‘I believe that the media-freedom act might influence the behaviour of [member states], including Hungary.’¹³⁴ The claims of concern about media monopolies

explicitly targeted countries like Poland and Hungary with conservative governments. Of course, this was Poland before arch-Eurocrat Donald Tusk of the Civic Platform replaced the conservative populist Mateusz Morawiecki as Prime Minister following the 2023 Polish parliamentary elections.

However, Brussels is relaxed about media monopolies in other EU member states. According to the Media Pluralism Monitor (MPM),¹³⁵ no European country is free from threats to media pluralism. The *Liberties Media Freedom Report 2023*, produced by the Civil Liberties Union for Europe, discovered that ‘strong media-ownership concentration persists and poses a significant risk to media pluralism’ throughout the EU.¹³⁶

And yet, Poland and Hungary were continually singled out by Brussels as posing a particular threat to media freedom.

In other areas of media freedom, the Civil Liberties Union for Europe reports significant threats across Europe – many of which were not to be found in Hungary or Poland at the time. For example, countries where journalists are most at risk of physical and verbal attacks included Belgium, Croatia, France, Germany, the Netherlands and Ireland, but not Poland and Hungary. The report also suggests that constraints on freedom of information are an EU-wide problem. Similarly, abusive lawsuits designed to silence critical journalism – so-called strategic lawsuits against public participation (SLAPPs) – are common across EU member states.

Poland and Hungary are not the only member states said to have a media-freedom problem. But they represented the standout villains for one simple reason: the political bent of the media in these countries tended to be nationalist, conservative and decidedly anti-woke. Such governments

Hungary's determination to provide an alternative to the cultural values of the EU elites has provoked the wrath of the EU and its partners

are a threat to the EU's cosmopolitan federalism because they stand up for national sovereignty and do not accept the woke social agenda. Hungary's determination to provide an alternative to the cultural values of the EU elites has

provoked the wrath of the EU and its partners, who are terrified other nations might follow suit.

In the name of freedom, Brussels intends to shackle media freedom. A new European Board for Media Services,¹³⁷ which replaces the European Regulators Group for Audiovisual Media Services (ERGA), now gives the EU a direct say over member states' media landscapes. It is an undisguised

drive to centralise media control away from national governments towards Brussels.

The EMFA is an attempt to control the media and politics of member states. It is driven by the same authoritarian dynamic underpinning the hate speech crusade: a defensive and almost paranoid attempt to outlaw anything that poses a threat to their project and thus exposes the crisis of authority at the heart of the EU.

The idea that the EU in any way favours media freedom, indeed free speech, is simply ludicrous. If they were genuinely concerned about media freedom and the fight against disinformation, the Brussels elites should surely have chastised *Politico* for its report on Donald Trump's recent speech about the car industry in America. Trumped promised voters that he would 'put a 100 per cent tariff on every single [foreign-made] car'. He then predicted that if he was not elected, there was 'going to be a bloodbath for the whole country; that's going to be the least of it.' *Politico* deliberately distorted Trump's figurative warnings about the future of US industry into a literal sensationalised headline which read: 'Trump says country faces bloodbath

The last thing the EU technocrats want is media pluralism. Their quest is to concentrate power in Brussels

if Biden wins in November'.¹³⁸

The EU's response? Nothing. Because *Politico* supports the Brussels bureaucracy, this deliberate misinterpretation

of Trump's speech was used to sensationalise a dire warning of a pending civil war in America and the supposed global threat posed by the 'far right'. It might be hard to think of a clearer case of 'disinformation'.

The last thing the EU technocrats want is media pluralism. Their quest is to concentrate power in Brussels. And freedom? Their modus operandi

is only the defence of their freedom to determine the narrative, to dictate what can and cannot be said and thus to protect their freedom to impose their technocratic values from above.

7.2 The EU's hate campaign against the peoples of Europe

Throughout this report, we have highlighted the systemic character of the censorship dynamic in the EU. Through laws like the DSA, the EMFA, the AIA, the Child Sexual Abuse Regulation and the regulation of political advertising,¹³⁹ Brussels has erected a legal censorship citadel aimed at curbing free speech to control the EU narrative. They constantly scan the horizon to spot and close any loophole through which opposing narratives might spring. It is a never-ending labour which is systematically throttling free speech in Europe.

But this never-ending quest is a fool's paradise. It can never end because it is a moral and political crusade that cannot achieve what it aims to do. It is impossible to legislate ideas, thoughts, and meaning – all expressed through language and speech – out of existence. The more you attempt to do this, the more you highlight the absence of freedom. Negative authority provides no positive impulse to inspire allegiance, only opposition. Unaccountable authority is authoritarianism, which only breeds resistance.

The hate speech narrative, with its fundamentally flawed assertion of a link between speech and violence, is not a truth beyond contestation. It is a ruse, not reality. It is the creation of a fragile technocratic oligarchy that fears free speech lest it question its version of the truth and raise fundamental questions about its right to rule. Most importantly, it is a means to one end: muzzling the people it regards with contempt, who are seen as targets for intervention rather than morally autonomous agents who can distinguish truth from fiction and from officially sanctioned disinformation.

The EU's narrative about curbing hate speech is essentially a hate campaign aimed at the peoples of Europe – one disguised as a defence of European values and democracy. Social media targeting seeks to muzzle both the 'social' and the media. Brussels' contempt for the ordinary people of Europe and their assumed inability to think for themselves means that behind this system of control lies a draconian authoritarianism. Free speech is the enemy: freedom from speech is its default response.

7.3 So, what is to be done?

What's needed now is a robust and public defence of free speech for all, no matter whether Brussels slurs them as 'far right', 'Putin's propagandists' or whatever else. The peoples of Europe must have the freedom to hear all of the arguments and judge for themselves.

More speech in the court of public opinion is the only long-term defence of democracy in Europe

The malicious and hateful prejudice of the EU elites, that ordinary people are too ignorant, stupid and prone to manipulation by demagogues, needs to be forcefully countered. We need a public campaign that holds the Brussels technocrats to account. We must bring like-minded free speech advocates together to shine the light of public opinion on the unholy alliance of unaccountables.

We need to flag their unaccountable flaggers, check their self-appointed checkers, and bring into the open their behind-closed-doors crusade to rob Europe of the foundations of democracy.

During the forthcoming EU elections, the goal should be to expose any and every attempt to muzzle views and speech deemed out of order by Brussels and their Big Tech minions. And just as the EU technocrats'

goal of legislating against free speech goes far beyond the election, we must also use the election to establish an ongoing campaign to defend free speech.

We need a campaign that argues that the only legitimate counter to hateful speech is free speech and that more speech in the court of public opinion is the only long-term defence of democracy in Europe.

8 Endnotes

- 1 eeas.europa.eu
- 2 weforum.org
- 3 Gold, A. & Fischer, S. (2023, 21 February). Chatbots trigger next misinformation nightmare. *Axios*. axios.com
- 4 Metz, C. (2023, May 1). ‘The Godfather of A.I.’ leaves Google and warns of danger ahead. *The New York Times*. nytimes.com
- 5 Scott, L. (2023, September 5). World faces ‘tech-enabled armageddon,’ Maria Ressa says. *Voice of America*. voanews.com
- 6 ‘Brussels Playbook’, *Politico* newsletter, 28 February 2024.
- 7 See, Foreign Electoral Interference Affecting EU Democratic Processes, Authority for European Political Parties and European Political Foundations - November 2023.
- 8 home-affairs.ec.europa.eu
- 9 See patrick-breyer.de
- 10 ‘Is this the most criticised draft EU law of all time?’ edri.org
- 11 edri.org
- 12 netzpolitik.org
- 13 ecpat.org
- 14 savanta.com
- 15 novus.se
- 16 eu-digital-services-act.com
- 17 gdpr-info.eu
- 18 ‘EU Commission faces complaint over microtargeted adverts on child sexual abuse bill’, *Euronews*, 16 November 2023. euronews.com
- 19 ‘EU Commission Accused of Using Illegal Micro-Targeted Ads’, the *European Conservative*, 17 November, 2023. europeanconservative.com
- 20 eu-digital-services-act.com See theguardian.com See ec.europa.eu
- 21 swp.de
- 22 See, “‘Pimmelgate’ – chronology of a debacle’, *Spiegel Panorama*, spiegel.de
- 23 semanticscholar.org
- 24 google.co.uk
- 25 politico.com
- 26 br.de
- 27 See stmi.bayern.de
- 28 fra.europa.eu Interesting to note that its Management Board consists of ‘independent’ experts, one appointed by each Member State, two European Commission representatives and one independent expert appointed by the Council of Europe
- 29 fra.europa.eu
- 30 See ‘Encouraging hate crime reporting: the role of Law Enforcement and other Authorities’, 2021. fra.europa.eu
- 31 fra.europa.eu
- 32 coe.int
- 33 rm.coe.int
- 34 coe.int
- 35 See the European Commission communication to the European Parliament and the Council, ‘A more inclusive and protective Europe: extending the list of EU crimes to hate speech and hate crime’ December 2021. eur-lex.europa.eu
- 36 See ‘The Brussels-led Pathologization of Speech’, *European Conservative*, 17 January 2024. europeanconservative.com
- 37 Allport, G. (1954) *The Nature of Prejudice*, Addison Wesley Publishing Company: Reading, MA. p51
- 38 ‘Factsheet - Hate speech’, Council of Europe, February 2012, p1
- 39 ‘Countering Online Hate Speech’, UNESCO, 2015, p8
- 40 ‘Factsheet - Hate Speech’, Council of Europe, November 2008, p2
- 41 ‘Hate Speech and Hate Crimes against LGBT Persons’, *Fundamental Rights Agency*, 2009, p1
- 42 ‘Homophobia and Discrimination on Grounds of Sexual Orientation and Gender Identity in the EU Member States Part II – The Social Situation’, *Fundamental Rights Agency*, 2009, p46
- 43 ‘Homophobia, transphobia and discrimination on grounds of sexual orientation and gender identity’, *Fundamental Rights Agency*, 2010, pp36–37
- 44 osce.org
- 45 Agnes Callamard, ‘Comments and Recommendation on ECRI General Policy Recommendation No. 15 on Combating Hate Speech’ (*Global Freedom of Expression at Columbia*, 20 May 2015), p4
- 46 Coleman, P(2016) *Censored*, Kairos Publications, Vienna, Second Edition. Kindle edition.
- 47 The drafting of the Universal Declaration of Human Rights began on 10 December 1948 and went through seven stages of drafting spanning over two years involving the active participation of nearly 50 countries
- 48 Jacob Mchangama (2011) ‘The Sordid Origin of Hate-Speech Laws’, *Hoover Institution, Policy Review*, No. 170, p1
- 49 Johannes Morsink (1999) *The Universal Declaration of Human Rights: Origins, Drafting, and Intent* (University of Pennsylvania Press), p66
- 50 Belgian representative, Fernand Dehousse A/C.3/SR.128, 9 November 1948, p414

- 51 Coleman, P (2016) *Censored: How European “Hate Speech” Laws are Threatening Freedom of Speech* (Kairos Publications). Kindle Edition
- 52 Coleman Ibid, A/PV.180, 9 December 1948, p54
- 53 E/CN.4/AC.2/SR/9, 10 December 1947, pp6–7
- 54 A/C.3/SR.1079, 20 October 1961, §9; E/CN.4/SR.377, 16 October 1953, §26
- 55 Eleanor Roosevelt, Commission on Human Rights, Sixth Session, E/CN.4/SR.174, 6 May 1950, p6
- 56 Sixteenth Session of the General Assembly, A/5000, 5 December 1961, p14
- 57 A series of high-profile anti-Semitic and racist manifestations in 1959-1960 provoked the ICERD. See Theodor Meron, ‘The Meaning and Reach of the International Convention on the Elimination of All Forms of Racial Discrimination’, (1985) 79 *American Journal of International Law* 283, p283
- 58 eur-lex.europa.eu
- 59 commission.europa.eu
- 60 European Union Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law, Article 1(2)
- 61 Article 1 of the Council Framework Decision 2008/913/JHA of 28 November 2008 on Combating Certain Forms and Expressions of Racism and Xenophobia states that: ‘Each Member State shall take the measure necessary to ensure that publicly inciting to violence or hatred directed against a group of persons or a member of such a group defined by reference to race, colour, religion, descent or national or ethnic origin’. eur-lex.europa.eu
- 62 European Commission (2020), *Combating racism and xenophobia: The Commission calls on Estonia and Romania to fully transpose EU law criminalising hate speech and hate crimes*, Infringement decisions, INF-20-1687, 30 October 2020
- 63 delano.lu
- 64 europarl.europa.eu
- 65 Council of Europe (2003b) *Explanatory Report, Additional Protocol to the Convention on cybercrime, Concerning the Criminalisation of Acts of a Racist and Xenophobic Nature Committed through Computer systems*, Strasbourg: CoE, Document STE No. 189 conventions.coe.int Cited in Pupavac, V (2012), *Language Rights From Free Speech to Linguistic Governance*, Palgrave MacMillan, Basingstoke and NY
- 66 ec.europa.eu
- 67 commission.europa.eu
- 68 transparency.fb.com
- 69 Google, ‘Transparency Report: YouTube Community Guidelines enforcement’ <transparencyreport.google.com/youtube-policy/removals> accessed 21 December 2022
- 70 ec.europa.eu
- 71 ec.europa.eu
- 72 ec.europa.eu
- 73 coe.int
- 74 ECRI General Policy Recommendation No. 7 on National Legislation to Combat Racism and Racial Discrimination, Council of Europe, adopted on 13 December 2002
- 75 Annual Report on ECRI’s Activities, 2013, p13
- 76 Agnes Callamard, ‘Comments and Recommendation No.15 on Combating Hate Speech’ (Global Freedom of Expression at Columbia, 20 May 2015), p4
- 77 coe.int
- 78 See, for example, the ECRI Report on Hungary, (fourth monitoring cycle), 24 February 2009, p13
- 79 See Rainbow Europe map and index 2022 (ilga-europe.org)
- 80 L. Walker, ‘Unacceptable violence’: Belgium to crack down on crimes against LGBTQ community, *The Brussels Times*, 18 May 2022
- 81 See, ‘Plan d’action fédéral 2021-2024, Pour une Belgique LGBTQI+ friendly, 17 May 2022
- 82 echr.coe.int
- 83 As initially developed in *Féret v Belgium* App no 15615/07 (ECtHR, 16 July 2009) and then in *Vejdeland v Sweden* App no 1813/07 (ECtHR, 9 May 2012) and *Atamanchuk v Russia* App no 4493/11 (ECtHR, 12 October 2020)
- 84 ec.europa.eu
- 85 See ilga-europe.org
- 86 hudoc.echr.coe.int
- 87 transparency.dsa.ec.europa.eu
- 88 See Facebook’s Transparency Centre: transparency.fb.com
- 89 ec.europa.eu
- 90 home-affairs.ec.europa.eu

ENDNOTES

- 91 europarl.europa.eu
- 92 ec.europa.eu
- 93 europarl.europa.eu
- 94 This is the point the European Network Against Racism (ENAR) makes in an intersectionality and hate crime briefing, ‘Hate crime provisions in EU Member States: the importance of an intersectional approach to ensure victims’, 2020 enar-eu.org
- 95 europarl.europa.eu
- 96 Simon, F. M., Altay, S., & Mercier, H. (2023). Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown. *Harvard Kennedy School (HKS) Misinformation Review*. doi.org
- 97 Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14). doi.org
- 98 Acerbi, A., Altay, S., & Mercier, H. (2022). Research note: Fighting misinformation or fighting for information? *Harvard Kennedy School (HKS) Misinformation Review*, 3(1). doi.org
- 99 Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425), 374–378. doi.org Interestingly, one study showed that during the Covid pandemic, a mere 12 people were behind 65 percent of online vaccine misinformation: counterhate.com
- 100 Altay, S., Berriche, M., Heuer, H., Farkas, J., & Rathje, S. (2023). A survey of expert views on misinformation: Definitions, determinants, solutions, and future of the field. *Harvard Kennedy School (HKS) Misinformation Review*, 4(4). doi.org
- 101 Altay, S., & Acerbi, A. (2023). People believe misinformation is a threat because they assume others are gullible. *New Media & Society*. doi.org
- 102 Mercier, H. (2020). *Not born yesterday: The science of who we trust and what we believe*. Princeton University Press.
- 103 cloudfront.net
- 104 the European Commission JRC Science for Policy Report 2020, p5
- 105 europarl.europa.eu
- 106 commission.europa.eu
- 107 MacAvaney S, Yao H-R, Yang E, Russell K, Goharian N, Frieder O (2019) Hate speech detection: Challenges and solutions. *PLoS ONE* 14(8): e0221152. doi.org
- 108 The authors cite the example of the following example taken from Facebook: ‘... The merciless Indian Savages, whose known rule of warfare, is an undistinguished destruction of all ages, sexes and conditions ...’. While this could be easily interpreted as hate speech, flagged and taken down, it is in fact, a direct quote from the US Declaration of Independence.
- 109 Natasha Duarte and Emma Llansó, ‘Mixed Messages? The Limits of Automated Social Media Content Analysis’ (CDT, 28 November 2017) cdt.org
- 110 fra.europa.eu
- 111 weaponizedword.org
- 112 Ibid, Natasha Duarte and Emma Llansó
- 113 See Börgmann, H (2023) ‘In the blind spot – Right-wing extremists on alternative and established platforms: Third conference report by ISD Germany on the state of research and options for action with regard to right-wing extremist radicalisation in digital spaces’, Institute for Strategic Dialogue, [ISD-Im_Toten_Winkel_3-EN-v3-digital.pdf](https://isd-im-toten-winkel_3-EN-v3-digital.pdf), p9
- 114 coe.int
- 115 opeuropa.eu
- 116 Seema Nagar, F. Barbhuiya, K. Dey (2023) ‘Towards more robust hate speech detection: using social context and user data’, *Social Network Analysis and Mining*, Springer, March 2023
- 117 Vosoughi, S., Roy, D., Aral, S., ‘The Spread of True and False News Online’, *Science*, Vol. 359, No. 6380, 9 March, 2018, pp1146–1151. Available at science.org
- 118 Lilian Edwards, *The Role and Responsibility of Internet Intermediaries in the Field of Copyright and Related Rights (WIPO 2016)*, 5
- 119 YU Wenguan, ‘Internet Intermediaries’ Liability for Online Illegal Hate Speech’ (2018) 13(3) *Frontiers of Law in China* 342, 356. Also see the Council of Europe leaflet: ‘The role of Internet intermediaries’. rm.coe.int
- 120 Jacob Mchangama, Natalie Alkiviadou and Raghav Mendiratta, ‘Rushing to Judgement: Are Short Mandatory Takedown Limits for Online Hate Speech Compatible with the Freedom of Expression?’ (*Justitia*, January 2021), 2 <futurefreespeech.com/wp-content/uploads/2021/01/FFS_Rushing-to-Judgment-3.pdf> accessed 21 December 2022.
- 121 See ‘technology Industry to Combat Deceptive use of AY in 2024 Elections; aielectionaccord.com
- 122 cnam.ie

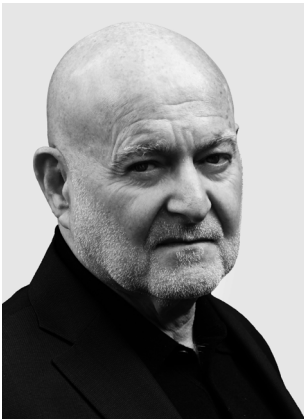
- 123 See Thomas O'Reilly, 'Small Irish Regulatory Authority To Enforce EU's Digital Services Act for 450 Million Europeans', *European Conservative*, 19 February 2024 europeanconservative.com
- 124 See 'Fact-Checking Meta's Fact-Checkers', Tamas Orban, *European Conservative*, 28 February 2024 europeanconservative.com
- 125 Quoted in *Politico's* 'Digital Bridge Newsletter', 23 February 2024
- 126 euvsdisinfo.eu
- 127 EUvsDisinfo newsletter, 1 February 2024.
- 128 economist.com
- 129 See 'The European Commission Recommends Big Tech Measures to Mitigate "Systemic Risks" for Electoral Processes', verfassungsblog.de
- 130 state.gov
- 131 ilga-europe.org
- 132 commission.europa.eu
- 133 See 'How the European Media Freedom Act could affect Hungary and Poland', ipi.media
- 134 theguardian.com
- 135 cmpf.eui.eu
- 136 liberties.eu
- 137 ec.europa.eu
- 138 politico.com
- 139 ec.europa.eu
- 140 Vosoughi, S., Roy, D., Aral, S., 'The Spread of True and False News Online', *Science*, Vol. 359, No. 6380, 9 March, 2018, pp1146–1151. Available at science.org

9 About the author

Dr Norman Lewis

Visiting Research Fellow

Managing Director, Futures-Diagnosis Ltd



Dr Norman Lewis is a writer, speaker and consultant on innovation and technology. He is recognised worldwide as an expert on future trends and user behaviours regarding technology innovation and adoption. Norman was formerly a Director at PwC, responsible for running its crowdsourced innovation programme. Prior to this, he was the Director of Technology Research at Orange. He is an Advisory Board Member of Bubbletone Blockchain in Telecom – the world's

first decentralised mobile roaming service.

He was also an Executive Board member of the MIT Communications Futures Programme and a former chairman of the ITU TELECOM Forum Programme Committee. He is a co-author of *Big Potatoes: the London manifesto for innovation*.

About MCC Brussels

At a time of unprecedented political polarisation, MCC Brussels is committed to providing a home for genuine policy deliberation and an in-depth exploration of the issues of our time.

MCC Brussels is committed to asking the hard questions and working with people of goodwill from all persuasions to find solutions to our most pressing problems. An initiative of MCC (Mathias Corvinus Collegium), the leading Hungarian educational forum, MCC Brussels was founded in the autumn of 2022 to make a case for celebrating true diversity of thought, diversity of views, and the diversity of European cultures and their values.

For some time, the EU has been assembling an incredible series of powers to control what can and cannot be said online.

At the heart of this is an unprecedented network of EU institutions, unaccountable NGOs, and Big Tech companies. Through agreements, laws, regulations and back-channel arrangements, these organisations tightly control online speech. This is a clear attempt by EU elites to determine what Europe's 448 million people can and cannot say online. Further sweeping regulations on online speech are planned.

The justification they give is the need to protect European democracy from hate speech and misinformation. But behind these invocations of democracy in fact lies a profoundly anti-democratic attitude towards European citizens.

Rather than Europe being under attack from "hate speech", European citizens are under attack from the hateful attitude of EU elites. The powers-that-be look down on European citizens as infants easily susceptible to manipulation who need to be insulated from harmful speech and ideas.

The EU elite are set to intensify their anti-speech efforts in coming months and years – especially in the wake of increasing challenges to their precious orthodoxies on the Green Deal, immigration, and EU integration. They are desperate to control the narrative and see controlling online speech – under the guise of combatting hate speech and disinformation – as central to this.

This report aims to challenge the Brussels 'hate speech' narrative, and return free speech to centre stage as the foundation for democracy in Europe.