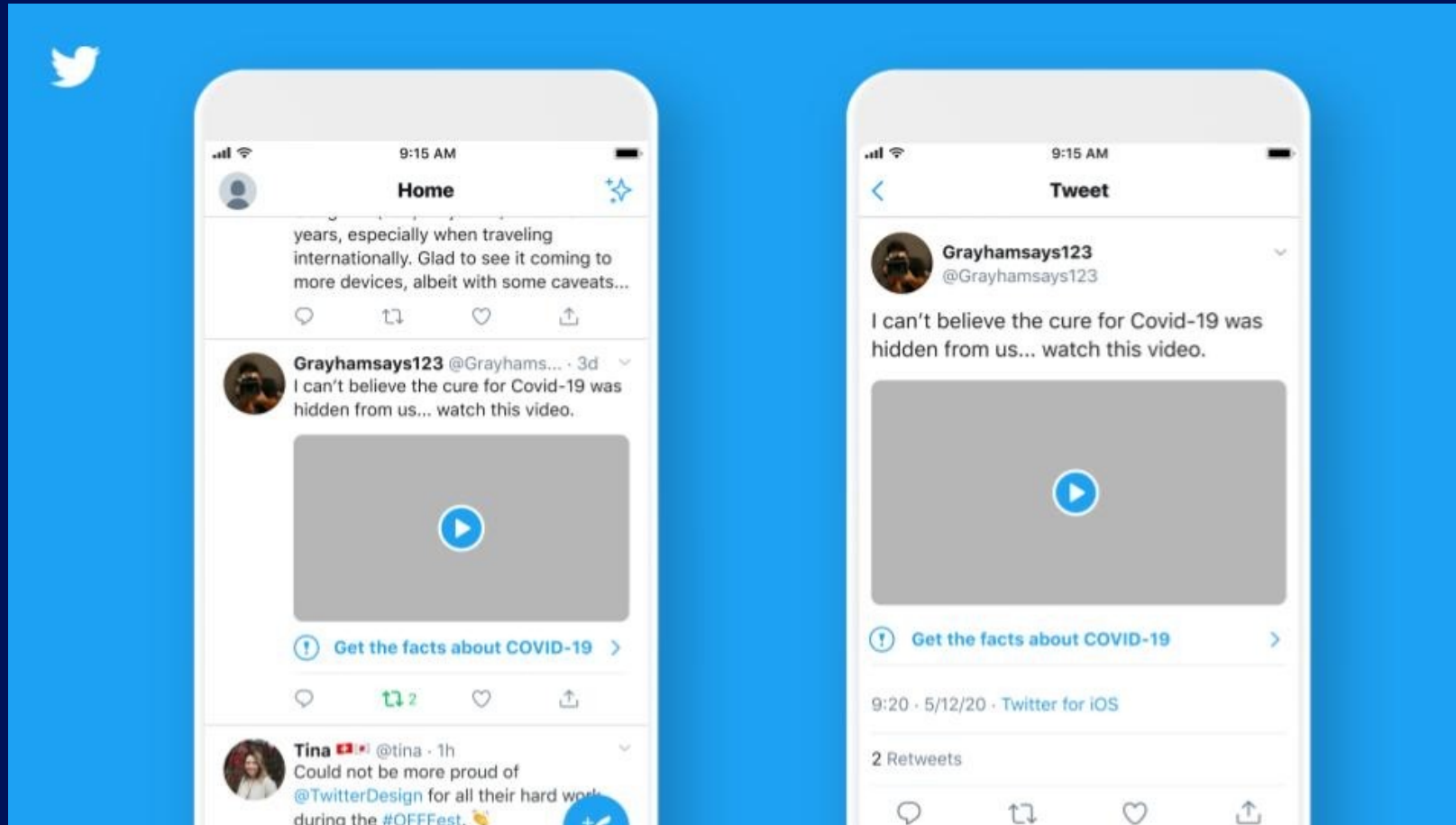


Disinformation and the law

Conflict Resolution Seminar @Leiden
Michael Klos LLM BSc | m.klos@law.leidenuniv.nl | 30 June 2022

Labelling



2018 Code of Practice on Disinformation

As provided under the Commission's Communication, for the purpose of this Code, the Commission as well as the High Level Expert Group in its report define "Disinformation" as "verifiably false or misleading information" which, cumulatively,

(a) "Is created, presented and disseminated for economic gain or to intentionally deceive the public"; and

(b) "May cause public harm", intended as "threats to democratic political and policymaking processes as well as public goods such as the protection of EU citizens' health, the environment or security".^[5]

The notion of "Disinformation" does not include misleading advertising, reporting errors, satire and parody, or clearly identified partisan news and commentary, and is without prejudice to binding legal obligations, self-regulatory advertising codes, and standards regarding misleading advertising.

Content Moderation Remedies (2021)

- Content Regulation
- Account Regulation
- Visibility Reductions
- Monetary actions
- Other

Content Moderation Remedies

- Is labelling harmless?
- Is removal not firing a canon on a mosquito?
- What effect has demonetisation?

What remedy fits the “crime” of disinformation?

Article 14 of the e-Commerce Directive

1. Where an information society service is provided that consists of the storage of information provided by a recipient of the service, Member States shall ensure that the service provider is not liable for the information stored at the request of a recipient of the service, on condition that:
 - a. the provider does not have actual knowledge of illegal activity or information and, as regards claims for damages, is not aware of facts or circumstances from which the illegal activity or information is apparent; or
 - b. the provider, upon obtaining such knowledge or awareness, acts expeditiously to remove or to disable access to the information.

Section 230(c) of the CDA 1996

Protection for “Good Samaritan” blocking and screening of offensive material

(1) Treatment of publisher or speaker

No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.

(2) Civil liability

No provider or user of an interactive computer service shall be held liable on account of—

(A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or [...]

Article 15 of the e-Commerce Directive

No general obligation to monitor

1. Member States shall not impose a general obligation on providers, when providing the services covered by Articles 12, 13 and 14, to monitor the information which they transmit or store, nor a general obligation actively to seek facts or circumstances indicating illegal activity.

To recall: article 14 shields service providers for liability for user-provided information as long as they have no knowledge or when they remove the information in question upon obtaining such knowledge.

JOIN/2020/8 final

“The COVID-19 ‘infodemic’ has demanded a rapid response from the EU and its Member States. Disinformation can have severe consequences: it can lead people to ignore official health advice and engage in risky behaviour, or have a negative impact on our democratic institutions, societies, as well as on our economic and financial situation. [...] Combatting the flow of disinformation, misinformation and foreign influence operations, including through proactive and positive communication, calls for action [...]. This work must be done in full respect of freedom of expression and other fundamental rights and democratic values.”

Community Guidelines

- **The existence or severity of COVID-19.** Acknowledging the existence and understanding the severity of COVID-19 is foundational to keeping people safe and aware of the dangers of this public health emergency. We remove claims that deny the existence of the disease or undermine the severity of COVID-19. This includes:
 - Claims that deny the existence of the COVID-19 disease or pandemic (however, we allow discussion or debate regarding whether the COVID-19 pandemic is transitioning to a less severe state, such as an endemic disease)
 - Claims that downplay the severity of COVID-19, such as:
 - In the context of discouraging vaccination or questioning the efficacy of vaccines, claims that COVID-19 is no more dangerous to people than the common flu or cold.
 - Claims that no one has died from COVID-19
 - Claims that having a flu jab or flu vaccine is more likely to kill you than COVID-19
 - Claims that the number of deaths caused by COVID-19 are much lower than the official figure (requires additional information and/or context)

Do not share false or misleading content: Do not share content in a way that you know is, or think may be, misleading or inaccurate, including misinformation or disinformation. Do not share content to interfere with or improperly influence an election or other civic process. We may prevent you from posting content from sites that are known to produce or contain misinformation. **Do not share content that directly contradicts guidance from leading global health organizations and public health authorities; including false information about the safety or efficacy of COVID-19 vaccines.** Do not share false content or information, including news stories, as though they are true or likely true. Do not post “deepfake” images or videos of others or otherwise post content that has been manipulated to deceive. Do not share content or endorse someone or something in exchange for personal benefit (including personal or family relationships, monetary payment, free products or services, or other value), unless you have included a clear and conspicuous notice of the personal benefit you receive and have otherwise complied with our [Advertising Policies](#).

2018 Code of Practice on Disinformation

II.D. Empowering consumers

Whereas:

- Consistently with Article 10 of the European Convention on Human Rights^[13] and the principle of freedom of opinion, Signatories should not be compelled by governments, nor should they adopt voluntary policies, to delete or prevent access to otherwise lawful content or messages solely on the basis that they are thought to be "false".

Indirect horizontal effect?

“Freedom of expression is a core value in our democracy and everyone must be able to express their opinion in accordance with the law. Certain fundamental rights, such as freedom of expression, are playing an increasingly important role in relations between citizens and businesses. Courts can give fundamental rights horizontal effect, for example by including fundamental rights protection in the interpretation of open private law norms. This assessment is made by a judge on a case-by-case basis.”

Aanhangsel Handelingen II 2019/20, nr. 3965, p. 13.

Open norms

Reasonableness and fairness (6:248 BW)

- A contract clause may not be applicable when this would be unacceptable to standards of reasonableness and fairness.

Unlawful act (6:162 BW)

- Infringement of a right
- Violation of a legal obligation
- Act or omission contrary to societal norms according to unwritten law

Three interim relief cases

Rechtbank Amsterdam (vzr.) 9 September 2020,
ECLI:NL:RBAMS:2020:4435 (*YouTube*):

- Removal of video's as Covid-19 'disinformation' because a medical doctor claimed that HCQ is an effective medicine against Covid-19 and social distancing does not work.
- Interesting because the plaintiff appeared in the video's is not the accountholder (no contract, did not agree on disinformation rules)

Rechtbank Amsterdam (vzr.) 13 October 2020,
ECLI:NL:RBAMS:2020:4966 (*Facebook*).

- Removal of a Facebook page/group/users/something (heh?) as disinformation

Three interim relief cases

Rb. Noord-Holland (vzr.) 6 October 2021,
ECLI:NL:RBNHO:2021:8539 (*LinkedIn*):

- Member of Dutch Parliament
- Multiple (+/- 12 posts) considered disinformation by LinkedIn
- Account restrictions

Disclaimer

Translations of Dutch legislation as well as Dutch case law are made only for this presentation and may not be 100% accurate.

Three questions

Concerning:

- Community guidelines
- Application/interpretation of the community guidelines
- The imposed remedy

Rechtbank Amsterdam (vzr.) 9 September 2020,
ECLI:NL:RBAMS:2020:4435 (*YouTube*).

“A doctor who claims, without conclusive evidence and scientifically-based tests, that HCQ or an alternative substance - which can be obtained without a prescription - works, misrepresents the public. In addition, it can be harmful and dangerous.

It may be the case that [plaintiff 2] believes that it has worked on 10 of his patients, that he believes in the drug (if it is used at an early stage) and that he wishes to have a public debate about it, but for that purpose he should also formulate his claims with this nuance. He does not do so.” (4.15)

Rechtbank Amsterdam (vzr.) 9 September 2020, ECLI:NL:RBAMS:2020:4435 (*YouTube*).

“As he said in the interviews, it is not part of the debate, which YouTube would have allowed, but his statements contain inaccurate information that is potentially harmful and dangerous. It is precisely as a doctor that he should realise this. Incidentally, this does not apply to the critical statements about the 1.5-metre measure and self-isolation [...] which may, however, contribute to debate.” (4.15)

Appleby v. U.K. (2003)

47. That provision, notwithstanding the acknowledged importance of freedom of expression, does not bestow any freedom of forum for the exercise of that right. While it is true that demographic, social, economic and technological developments are changing the ways in which people move around and come into contact with each other, the Court is not persuaded that this requires the automatic creation of rights of entry to private property, or even, necessarily, to all publicly owned property (government offices and ministries, for instance). Where, however, the bar on access to property has the effect of preventing any effective exercise of freedom of expression or it can be said that the essence of the right has been destroyed, the Court would not exclude that a positive obligation could arise for the State to protect the enjoyment of the Convention rights by regulating property rights. A corporate town where the entire municipality is controlled by a private body might be an example (see *Marsh v. Alabama*, cited at paragraph 26 above).

Rechtbank Amsterdam (vzr.) 13 October 2020,
ECLI:NL:RBAMS:2020:4966 (*Facebook*).

“The mere fact that Facebook is a powerful party and its platform has an enormous reach is not enough to deviate from that principle. It follows from ECtHR case law that it is for the State to take appropriate measures if the right to freedom of expression cannot be exercised effectively because of the dominance of powerful private parties. Positive obligations which may arise as a result are, in the given circumstances, at present non-existent.” (4.24)

LinkedIn and the DSA

“The platforms, including LinkedIn, are formally private parties. However, in their content moderation they do comply with requests from the (European) government, namely the call to do more to prevent disinformation about Covid-19, and the subsequent communication, in which explicit reference is made to the Digital Services Act as forthcoming legislation.”

“As said, in Europe a proposal for the Digital Services Act (hereafter DSA) is pending (see 2.8). This proposal is therefore not yet a European regulation, but LinkedIn has invoked it. It has pointed out that the proposal shows that a social media platform is not obliged to allow every (legal) content offered on its platform (no must-carry obligation).”

(Rb. Noord-Holland (vzr.) 6 October 2021, ECLI:NL:RBNHO:2021:8539 (*LinkedIn*), r.o. 4.5 and 4.13)

Relevant articles

- Article 12: Terms and conditions
- Article 15: Statement of reasons (removal / disable access)

Article 2(p): ‘content moderation’ means the activities undertaken by providers of intermediary services aimed at detecting, identifying and addressing illegal content or information incompatible with their terms and conditions, provided by recipients of the service, including measures taken that affect the availability, visibility and accessibility of that illegal content or that information, such as demotion, disabling of access to, or removal thereof, or the recipients’ ability to provide that information, such as the termination or suspension of a recipient’s account;

Be Trustworthy

We require you to use your true identity on LinkedIn, provide accurate information about yourself or your organization, and only share information that is real and authentic.



Do not share false or misleading content: Do not share content in a way that you know is, or think may be, misleading or inaccurate, including misinformation or disinformation. Do not share content to interfere with or improperly influence an election or another civic process. We may prevent you from posting content from sites that are known to produce or contain misinformation. Do not share content that directly contradicts guidance from leading global health organizations and public health authorities. Do not post content that denies a well-documented historical event such as the Holocaust or slavery. Do not share false content or information, including news stories, that present untrue facts or events as though they are true or likely true. Do not post “deepfake” images or videos of others or otherwise post content that has been manipulated to deceive. Do not share content or endorse someone or something in exchange for personal benefit (including personal or family relationships, monetary payment, free products or services, or other value), unless you have included a clear and conspicuous notice of the personal benefit you receive and have otherwise complied with our [Advertising Policies](#).

Statement of reasons?

“Seen in this light, the restriction is too severe. There is no clear policy, there has been little or no communication and, in so far as there has been any communication, it contains no reasoning beyond a mere reference to the Licence Agreement or the Policy Statement. The explanation given by LinkedIn on 19 June 2021 in the e-mail (2.16) is much too late and inadequate and [Member of Parliament]’s request for communication was not properly addressed.”

(Rb. Noord-Holland (vzr.) 6 October 2021, ECLI:NL:RBNHO:2021:8539 (*LinkedIn*), r.o. 4.22)

Statement of reasons?

“LinkedIn should take into account that its duty to give reasons in concrete cases is greater the more vaguely the policy that it claims to be applying with its decisions is described and/or is less recognisable. The current description is insufficiently informative. The further explanation LinkedIn has given at the hearing (critical content is allowed but no medical claims without scientific foundation) helps, but leaves room for considerable clarification.”

(Rb. Noord-Holland (vzr.) 6 October 2021, ECLI:NL:RBNHO:2021:8539 (*LinkedIn*), r.o. 4.22)

“Moderation is hard”

“LinkedIn has argued that content moderation is complicated. That is all the more reason to attach great importance to the use of clearly written policies and a procedure that pays attention to clear communication with the user and a good exchange of views. In this way, both sides can learn and the possibility of this learning process - also in the case of stubborn users - can provide sufficient justification for a sanctions policy that may ultimately result in the banishment from the platform.”

(Rb. Noord-Holland (vzr.) 6 October 2021, ECLI:NL:RBNHO:2021:8539 (*LinkedIn*), r.o. 4.23)

Conclusion?

YouTube: Indirect horizontal application of freedom of expression rights. A large portion of the content qualified as disinformation under the Code of Practice and the Communication. This is different for claims made about social distancing and self-isolation. What would the outcome be without the Code of Practice?

Facebook: No indirect horizontal application of freedom of expression rights. Facebook does not render freedom of expression rights meaningless. “Positive obligations which may arise as a result are, in the given circumstances, at present non-existent” (4.24)

Conclusion?

LinkedIn: on account restrictions: “Seen in this light, the restriction is too severe. There is no clear policy, there has been little or no communication and, in so far as there has been any communication, it contains no reasoning [...]. The current description is insufficiently informative. The further explanation LinkedIn has given at the hearing (critical content is allowed but no medical claims without scientific foundation) helps, but leaves room for considerable clarification.” (4.22)