

Digital Wildfire Delphi round 2

Page 1: Digital Wildfire Delphi round 2: Instructions

Dear Panellist,

This is the second of three Questionnaires for the Digital Wildfires Delphi (DWD). The purpose of a Delphi questionnaire is to stimulate discussion and deliberation amongst those with particular expertise about a policy issue such as the responsible uses of social media. The specific aim of the Digital Wildfires Delphi is to explore the technical and political feasibility of governing and regulating the responsible use of social media in England and Wales.

*A central issue is whether the potential harm caused by a digital wildfire outweighs the potential harm caused by the (temporary) removal of freedom of speech, where a 'digital wildfire' is being defined as, **'social media interactions in which misleading or provocative content spreads rapidly with very negative impact.'***

Thank you for taking the time to complete round 1 of the Delphi. Round 2 seeks to:

- confirm panellists' views on the usefulness of the concept of 'digital wildfires' in debates over the responsible governance of social media communications in the light of responses to Round 1;
- establish whether their views have altered as a consequence of other panellists' views on responsible governance; and, finally, to
- provide respondents with an opportunity to discuss the technical and political feasibility of different methods for governing harmful social media communications.

For round 2 you are asked to:

1. Read the summary report of the round 1 responses, sent to you via email and also available [here](#).
2. Answer the round 2 questions by checking the appropriate boxes and adding comments in the text boxes. Feel free to write as much as you would like, suggest clarifications, argue in favour of or against issues, and pose further questions for consideration in the third questionnaire.

Page 2: ID and Question 2.1 The usefulness of the concept of digital wildfire

Please write your name here. (This is so that we can keep track of who has responded to the questionnaire. Your name will only be seen by the project team and will not be included in any public reports about the project).

If you would rather not give your name, please indicate whether you are a:

Please select exactly 1 answer(s).

- ☐ lawyer
- ☐ employee at a social media platform
- ☐ employee at an institution (for example: educational establishment, school, government agency)
- ☐ none of the above and responding as a social media user

The usefulness of the concept of 'digital wildfire' Please indicate the strength of your agreement or disagreement with the following statements and, using the free text box, give your reasons for this choice.

	Strongly agree	Agree	Disagree	Strongly disagree	Reasons
Q2.1.1. 'Digital wildfire' is an ambiguous term that confuses more than it enlightens policy debates over the regulation of harmful social media communications.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="text"/>
Q2.1.2. For the purposes of regulating harmful social media communications, the term 'digital wildfire' is better replaced with references to specific offences (such as 'defamation', 'incitement', 'libel', 'menacing' and 'obscenity').	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="text"/>
Q2.1.3. Social media communications should only be regarded as harmful if they can be unambiguously related to an existing offence (such as 'defamation', 'incitement', 'libel', 'menacing' and 'obscenity').	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="text"/>

Page 3: Question 2.2: The responsibility for regulating social media communications

Please indicate the strength of your agreement or disagreement with the following statements and, using the free text box, give your reasons for this choice. Please also use the 'Reasons' text box to identify any other individuals or organisations you think ought to be responsible for regulating harmful social media communications.

	Strongly agree	Agree	Disagree	Strongly disagree	Reasons
Q2.2.1. Those with responsibility for the enforcement of civil and criminal laws on harmful communications ought to be the only actors empowered to regulate communications through social media	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>
Q2.2.2. Institutions with the capacity to prevent harmful social media communications (such as families, schools, work places and leisure venues) ought to be statutorily obliged to promote the responsible uses of social media	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>
Q2.2.3. Platforms providing social media services ought to be obliged, under the European Union's Responsible Research and Innovation framework, to assess the potential harmful social impact of their services and to mitigate any identified harmful impact through further rounds of technological innovation, including revisions to the Terms and Conditions of using their services and/or specific provisions for young people such as the right to easily edit or delete their online content	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>
Q2.2.4. The users of social media platforms ought to have the primary responsibility for regulating their own and other users' harmful communications	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>

Page 4: Question 2.3: The technical and political feasibility of regulating social media communications

The following methods of regulating social media communications were identified in responses to the first round questionnaire. Please rate how feasible you think they are, using the attached rating scale for the technical and political feasibility of regulating social media communications (see Appendix One below). Using the free text box, please give your reasons for this rating and identify any other methods of regulation that you think have been missed in responses to the first round questionnaire. Please also use this free text box to note any unintended consequences of using these methods of regulation which are already known or can be plausibly anticipated.

	Technical Feasibility *					Reasons	Political Feasibility					Reasons
	1	2	3	4	5		1	2	3	4	5	
Q2.3.1. The criminal prosecution of those responsible for posting harmful social media communications.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>
Q2.3.2. The criminal prosecution of those responsible for propagating harmful social media communications (for example through retweets or mentions).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>
Q2.3.3. The prosecution under criminal law of platforms for the dereliction of their duty of care towards users of their services or 'host recklessness'.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>
Q2.3.4. The use of administrative penalties to fine platforms for failing to monitor and remove harmful social media communications.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<div></div>

Q2.3.5. The use of digital technology to disrupt the impact of harmful social media communications through, for example, software that controls access to certain parts of the internet and at certain times of the day for particular devices (such as tablets and smart phones owned by children).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Q2.3.6. The use of platform technology (such as 'panic buttons' and report buttons) to trigger intervention by platform administrators to censor harmful communications	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Q2.3.7. The use of platform technology to trigger intervention by platform administrators to issue cautions and warning notices to accounts posting/propagating harmful communications with the threat of closing accounts and reporting account holders to the authorities for criminal prosecution.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Q2.3.8 The legal right for children and young people to easily edit or delete all online content they have created	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Q2.3.9. Removing the license of platforms to offer services in a jurisdiction if they are found to have failed their duty of care to service users.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

Q2.3.10. 'Kite-marking' and officially endorsing those platforms that demonstrate a duty of care to the users of their services.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Q2.3.11. Taxing those platforms that demonstrate a duty of care to the users of their services at a lower rate than those failing to demonstrate a duty of care.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Q2.3.12. Triggering the self-regulation of social media users through educational programmes and public campaigns, including case studies of victimisation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Q2.3.13. Promote freedom of speech to enable and stimulate enlightening counter-speech to harmful posts and/or the shaming and ridicule of harmful posts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

Appendix 1:

Scale	Technical Feasibility	Political Feasibility
(1) = Definitely feasible	<ul style="list-style-type: none"> • Can be implemented • No further research and development required • Necessary resources (finance, labour etc.) are presently available 	<ul style="list-style-type: none"> • No major political obstacles • Will be acceptable to the general public

(2) = Probably feasible	<ul style="list-style-type: none"> • Some indication that this can be implemented • Some research and development still required • Available resources would have to be supplemented 	<ul style="list-style-type: none"> • Some minor political obstacles • Further consideration may need to be given to public reaction, although some indication exists that the proposed resolution of this issue may be acceptable
(3) = May or may not be feasible	<ul style="list-style-type: none"> • Contradictory evidence that this can be implemented • Indeterminable research and development effort needed (existing resources may be inadequate) • Increase in available resources needed 	<ul style="list-style-type: none"> • Political obstacles • Some indication that this may not be acceptable to a large proportion of the general public
(4) = Probably unfeasible	<ul style="list-style-type: none"> • Some indication that this cannot be implemented • Major research and development effort needed (existing resources are inadequate) • Large scale increase in resources needed 	<ul style="list-style-type: none"> • Major political obstacles • Not acceptable to a large proportion of the general public
(5) = Definitely unfeasible	<ul style="list-style-type: none"> • Cannot be implemented • Basic research needed • Unprecedented allocation of resources needed 	<ul style="list-style-type: none"> • Politically unacceptable • Completely unacceptable to the general public

Thank you!

Thank you for completing our questionnaire. If you have any queries or comments about the Delphi, please contact helena.webb@cs.ox.ac.uk.

If you would like more information about the Digital Wildfire project, please look at our website www.digitalwildfire.org and You Tube channel 'Digital Wildfire project'.