

# ICO call for views: Data protection and content moderation.

The ICO is calling for views to support the development of guidance that will provide clarity on how data protection law applies to online content moderation processes.

Content moderation is the process of analysing content generated by users of online platforms to assess whether it meets certain standards or whether it is appropriate for a given context. This content can often include personal data and the analysis may trigger subsequent action to be taken, such as removal of the content, escalation to human moderator or application of a content warning. We have a role to ensure that personal data is processed in compliance with data protection law during content moderation.

This work is the first step in meeting commitments set out in our November 2022 [Joint Statement with Ofcom](https://ico.org.uk/media/about-the-ico/documents/4022906/online-safety-and-data-protection-a-joint-statement-by-ofcom-and-the-ico.pdf) to provide certainty for organisations in scope of the future UK online safety regime

Our guidance, in parallel with our ongoing collaborative work with Ofcom, will support organisations to meet both their data protection and online safety responsibilities. It will also support other organisations developing and deploying content moderation in a non-online safety context.

As part of our call for views we want to understand more about:

* How content moderation solutions use people’s personal data, and how the solutions are being used or developed more generally
* Where clarification is needed on the application of UK GDPR, DPA 2018 and PECR to content moderation
* What challenges organisations are facing relating to data protection when they deploy content moderation processes

We recognise that content moderation is a developing area and new approaches are continuing to emerge. We are launching this call for views to develop our knowledge and to ensure that the guidance we provide about meeting data protection expectations is meaningful to organisations.

## Responding to the call for views

We want to hear from organisations and individuals with expertise, experience or an interest in the use of content moderation technology.

Please complete the call for views through our online survey, available through this link: [survey link](https://online1.snapsurveys.com/q3j3n4).

Alternatively you can download this Word document and email your response to [onlinesafetyteam@ico.org.uk](mailto:onlinesafetyteam@ico.org.uk).

You don’t need to answer every question - some of the questions may not be relevant to you or your organisation, so please skip these as necessary.

This call for views will remain open until **9 June 2023**.

If you have any general queries about the call for views or would like further information, please email us at [onlinesafetyteam@ico.org.uk](mailto:onlinesafetyteam@ico.org.uk).

## Privacy statement

For this consultation, we may publish a summary of the responses but will not publish the actual responses received from organisations or individuals. If we publish a summary of the responses, information and views will not be attributed to individual respondents.

Should we receive an FOI request for your response we will need to consider whether we make it available. However, at this point, we would always seek to consult with you for your views on the disclosure of this information before any decision is made.

For more information about what we do with personal data please see our [privacy notice](https://ico.org.uk/global/privacy-notice/responding-to-our-consultation-requests-and-surveys/).

Please note that we are using the platform Snap Surveys to gather this information. Any data collected by Snap Surveys for ICO is stored on UK servers. You can read their Privacy Policy at [www.snapsurveys.com/survey-software/privacy-policy-uk/](http://www.snapsurveys.com/survey-software/privacy-policy-uk/)

# Questions

## About you

1. Are you answering this call for views as:

A representative of an organisation that develops content moderation solutions for others to use

A representative of an organisation that uses content moderation solutions developed by a third-party

A representative of an organisation that develops content moderation solutions for use in-house

A representative of an organisation that develops content moderation solutions for use in-house AND solutions for others to use

Other (please specify in Q2)

1. If you selected ‘Other’ in the previous question, please pick one option:

☐ A representative of a professional, industry or trade association

☐ A representative of a third sector/civil society body (eg charity, voluntary and community organisation, social enterprise or think tank)

☐ A representative of a public body

☐ A representative of a private sector organisation

☐ An academic, academic research group or academic institution

☐ An individual acting in a private capacity (eg someone providing their views as a member of the public)

Other (please specify) …………………

1. What sector(s) is your organisation involved in? Tick all that apply. (This question is for developers and/or users of content moderation systems only)

Social media

Forums or chatrooms

Review sites

Blogging

Gaming

Retail

P2P marketplaces

Volunteering

Job searching

Search engines

Accommodation searching

Adult entertainment

Dating

Crowdfunding or fundraising websites

Content subscription or fansites

Private messaging

Video sharing or livestreaming

Video calling

Other (please specify):………….

1. For the content moderation solutions you develop for others to use, where are your clients based? Tick all that apply. (This question is for developers of content moderation systems for others to use only)

☐ UK

☐ Europe (excluding UK)

☐ Rest of the world (excluding UK and Europe)

1. How would you describe your organisation? (This question is for developers and users of content moderation systems only)

0 to 9 members of staff

10 to 249 members of staff

250 to 499 members of staff

500 or more members of staff

Not applicable or not sure

1. What is the name of your organisation and your role (if applicable)?

## Content moderation overview

We are asking the following questions to improve our understanding of how content moderation systems work and challenges to their development and deployment

1. Please tell us about the content moderation systems you have used, developed or experienced.

*We are interested in information including what technology is used, the purposes that content moderation serves, what areas of a service are moderated, the effectiveness of different approaches, and whether automation is combined with human moderation.*

|  |
| --- |
|  |

1. What do you consider to be the main challenges to development and deployment of content moderation solutions?

|  |
| --- |
|  |

## Data protection

We are asking these questions to improve our understanding about the data protection risks and challenges associated with content moderation, and any uncertainties that exist in this area.

1. What do you perceive to be the main data protection risks and challenges associated with the development and deployment of content moderation systems?

|  |
| --- |
|  |

1. Please list any particular areas of uncertainty where clarification is needed in the application of data protection law (the UK GDPR and the DPA 2018) and/or the Privacy and Electronic Communications Regulations to content moderation.

|  |
| --- |
|  |

## Information gathered and used

We are asking these questions to improve our understanding of what information is used in content moderation systems.

Please answer the following questions in relation to a content moderation system that you have used or developed most often or most recently.

The questions in this section are for developers and/or users of content moderation systems only.

1. Does the content moderation system you have used or developed require data sets for training and/or testing?

Yes

No

Unsure/ not applicable

If yes, please provide information about from where this data is sourced

|  |
| --- |
|  |

1. Is the content connected with an individual service user’s profile while it’s being moderated?

Yes

No

Unsure/ not applicable

Please provide more information

|  |
| --- |
|  |

1. Are there policies and/or governance mechanisms around human moderators accessing information about individual service users?

Yes

No

Unsure/ not applicable

If yes, please provide more information

|  |
| --- |
|  |

1. Do the answers you've provided in this section differ across different systems you have used and/or developed? If so, how? across different systems? If so, how?

|  |
| --- |
|  |

## Provision of information to users

We are asking these questions to improve our understanding of what information is provided to individuals about content moderation.

If you are answering these questions as a developer or user of content moderation solutions, please answer in relation to a moderation system that you have used or developed most often or most recently.

1. Are individual service users made aware that content moderation is being used?

Yes

No

Unsure/ not applicable

If yes, how is this information given to them?

|  |
| --- |
|  |

1. Is information given to individual users if action is taken on their content by a moderation system?

Yes

No

Unsure/ not applicable

If yes, what information is given to users?

|  |
| --- |
|  |

1. If you answered ‘yes’ in question 16, are there mechanisms in place for users to appeal against that action?

|  |
| --- |
|  |

## Reporting

We are asking these questions to improve our understanding about reporting mechanisms, including how information about individuals is used in content reporting.

If you are answering these questions as a developer or user of content moderation solutions, please answer in relation to a moderation system that you have used or developed most often or most recently.

1. Do mechanisms exist for individual service users to report content?

Yes

No

Unsure/ not applicable

If yes, how does this work and what action is taken when a report is made?

|  |
| --- |
|  |

1. If you answered ‘yes’ to question 18, to what extent is information about service users included in those reports?

|  |
| --- |
|  |

1. If illegal content is detected, is this subsequently reported to another organisation?

Yes

No

Unsure/ not applicable

If yes, what organisation is illegal content reported to and what information about individual service users is included in the reports?

|  |
| --- |
|  |

## Additional considerations

1. Please describe any further issues (not already covered) that you feel it would be beneficial for the ICO to consider in relation to content moderation.

|  |
| --- |
|  |

1. The ICO is planning to hold further engagement exercises as part of its programme of work on online safety technologies. Would you like to participate in future engagement activities?

Yes

No

If yes, please provide the best contact details:………………………………….

## Impacts on your organisation

We would like to understand more about how future ICO guidance and support in this area may impact you or your organisation, and issues that you would find it useful for guidance to cover.

The questions in this section are for developers and/or users of content moderation systems only.

1. Please tell us about the cost implications of content moderation approaches. What aspects of these costs are linked to complying with data protection and privacy law?

|  |
| --- |
|  |

1. Please provide information about any other impacts (positive or negative) you are likely to experience from applying data protection and privacy law to your content moderation approaches.

|  |
| --- |
|  |

1. Who in your organisation is likely to use guidance on content moderation? (please provide job titles or roles, not people’s names).

|  |
| --- |
|  |

1. If greater clarity were provided on how the data protection regime and the Privacy and Electronic Communications Regulations (PECR) apply to online content moderation processes, how might that benefit your organisation? Tick all that apply.

Confidence that you are providing a compliant service / product.

Marketing / promoting brand

Increased customer confidence / reassurance

Increased revenues or profits

Reduced legal or advisory costs

Would not benefit

Other (please specify): …………..

## Baseline Monitoring Information

We are interested in current levels of understanding of the UK data protection regime and the Privacy and Electronic Communications Regulations among developers and users of content moderation solutions, and how confident organisations feel in complying with these.

The questions in this section are for developers and/or users of content moderation systems only.

1. Please rate your general understanding of the UK data protection regime (UK General Data Protection Regulation and the Data Protection Act 2018).

1 – very low

2 – low

3 – neither low, nor high

4 – high

5 – very high

Unsure/don’t know

1. Please rate your general understanding of the Privacy and Electronic Communications Regulations (PECR).

1 – very low

2 – low

3 – neither low, nor high

4 – high

5 – very high

Unsure/don’t know

1. Please rate how confident you are in ensuring the content moderation systems you develop and/or use comply with data protection legislation. Please explain the reasoning for your choice.

1 – very low

2 – low

3 – neither low, nor high

4 – high

5 – very high

Unsure/don’t know

Please explain why:

|  |
| --- |
|  |

1. Please rate how confident you are in ensuring the content moderation systems you develop and/or deploy comply with PECR legislation. Please explain the reasoning for your choice.

1 – very low

2 – low

3 – neither low, nor high

4 – high

5 – very high

Unsure/don’t know

Please explain why:

|  |
| --- |
|  |