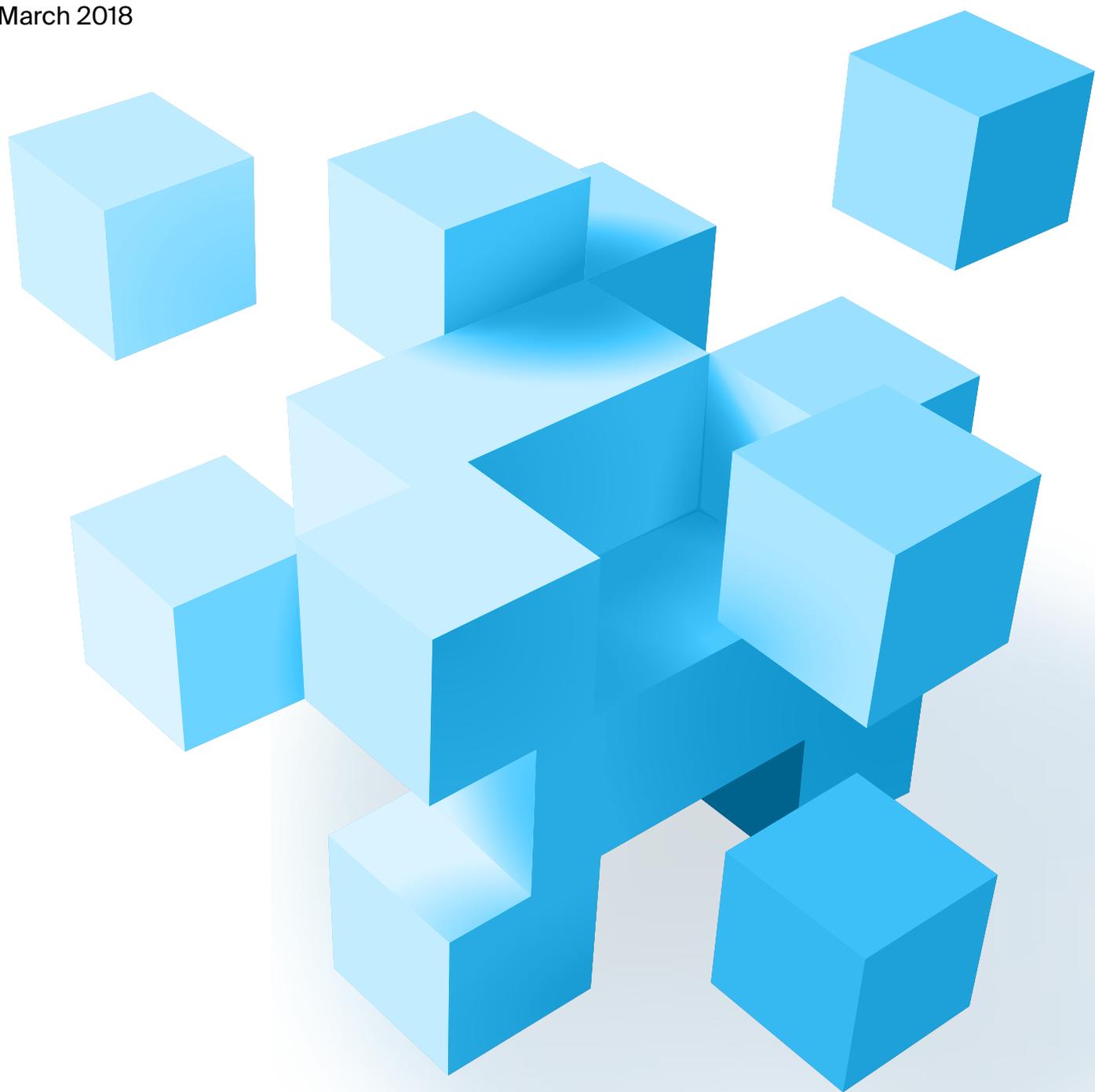


Challenging hate speech on social media platforms



A Guide for Advertisers

March 2018



INTRODUCTION

One of the more unpleasant consequences of the growth of social media has been the proliferation of hateful comments posted on social media platforms.

Social media has created an environment in which far too many people feel it is ok to make vile, threatening comments without compunction.

As the organisation representing leading advertisers, ISBA believes it is important to stand up to hate, to ensure that those who resort to hate speech understand that responsible brands will not tolerate intolerance.

We believe that turning a blind eye or deleting vile comments is not enough. In fact we think it risks making us complicit. So over the last few months we have been working with one of our members, Nationwide Building Society, to explore effective ways of tackling this form of abuse when it appears in social media feeds.

We have also collaborated with the Metropolitan Police and [Stop Hate UK](#) (one of the leading national organisations working to challenge all forms of hate crime and discrimination that has 'Trusted Flagging Status' with YouTube, Twitter, Instagram and Facebook) to develop this guidance which we hope all members will adopt.

GUIDANCE

Our aim has been to establish a clear, simple route to counter hate speech that does not inflame or open further dialogue with the perpetrator.

GUIDING PRINCIPLES

- **Don't be abusive yourself:** make sure your words and any content you share don't spread hate.
- **Don't spread hate:** avoid spreading hate, so if abuse has come from someone with only have a few followers don't give them any public attention – just report them to the relevant platform.
- **React promptly:** responding to old comments risks reviving a conversation and have an adverse effect.
- **Be consistent** in what and how you tackle hate speech.

IDENTIFY

Define what you regard as unacceptable.

Hate speech is typically directed towards another person or group on the grounds of race, nationality, gender, sexual orientation, colour, ethnic origin and religion with the intention to harass, distress, offend and/or threaten. In this particular context it is articulated in the form of comments made on advertiser content on social media platforms.

All the platforms have acceptable use policies which will cover these types of comments, although they differ from platform to platform.

It is important to note that not all hateful comments will meet the threshold set by the police or prosecuting authorities for criminal action. That does not mean that you need to set the same threshold; you can set your bar lower.

We would suggest your policy covers: threats of violence to a named individual and communications which may be considered grossly offensive (sexist, racist, homophobic, threatening etc), indecent, obscene or false. You may wish also to consider reference to violence which is not a real threat, which includes offensive language but doesn't target specific individuals.

You can find more information on what constitutes online hate crime [here](#).

FLAG

Flag comments which meet your criteria and have been posted with someone with a significant following by using **#challengehate** and **#reported**, to show the perpetrator and others that their comment is unacceptable.

If you also decide you would like to respond to their comments, choose your words carefully to avoid inflaming things.

- Don't get personal, think about your tone, express support for those under attack, reinforce notions of civility, respect and politeness, use language that will encourage support for your position and where appropriate use facts and statistics.

REPORT

There are two ways to report hate speech.

All hate speech should be reported to the host platform which you should do from your official account (not the individual account of a member of the team) via the relevant reporting tools:

- **Twitter:** report comments [here >](#)
- **Facebook:** report comments [here >](#)
- **YouTube:** report comments [here >](#)
- **Instagram:** report comments [here >](#)

If you think the comment may breach the law (eg there is a specific threat of violence aimed at an individual or it is grossly offensive) and you think it originates in the UK you should report it to the police.

This is easily done through the [True Vision online portal](#), a Government-funded web site designed for reporting online hate crime which will pass your report to your local police force.

MONITOR

Keep track of the volume of comments you flag and report.

ISBA is asking members to report this data to us on a quarterly basis so we can develop an overall picture of the scale of the issue and the effectiveness of this action.

SUPPORT

Where an individual or individuals to whom you owe a duty of care have been the victims of hateful comments, you should offer support either through your own employee care systems or via [Victim Support](#). [Stop Hate UK](#) is also a useful resource in this respect.

CONTACTS

For more information on any aspect of this guidance, please contact:

- Tanya Joseph, ISBA. [Contact >](#)
- DI John Donovan, Crime Policy, Hate Crime, Metropolitan Police. [Contact >](#)

ISBA represents the leading UK advertisers. We champion the needs of marketers through advocacy and offer our members thought leadership, consultancy, a programme of capability and networking.

We influence necessary change, speaking with one voice to all stakeholders including agencies, regulators, platform owners and government.

Our members represent over 3000 brands across a range of sectors.

For more information please visit www.isba.org.uk



ONE VOICE