Facebook wants your nude photos to prevent 'revenge porn' – here's why you should be sceptical

November 15 2017, by Amy Binns

Credit: CC0 Public Domain

Facebook's latest attempt to tackle the non-consensual sharing of sexual pictures (often known as "revenge porn") appeared so wrong-headed that

at first it seemed like a joke. But the social network has made clear its system of asking users to send in explicit images that they don't want to appear on the site is a real pilot programme being tested in Australia.

Facebook's motivation is right and proper: to help women (and men) worried that their ex-partners may shame or manipulate them by uploading sexual images taken during the relationship. This unwanted sharing can have devastating consequences. Even the threat that the images could be shared can be used by controlling, violent abusers to force their victims into line, as has been recognised by a new Scottish law to criminalise this.

To prevent people falling victim to this practice, users are urged to use Facebook's Messenger app to send themselves any pictures at risk of being shared. Facebook will then "hash" the image, creating a numerical fingerprint of it. The picture itself can then be deleted and Facebook has said images will not be stored permanently on their servers. When another Facebook user uploads a picture, it will be run through the database of hashes. If it matches an image in the database, it will be blocked and cannot be posted or shared on Facebook.

Would this work? If a picture uploaded by a vengeful ex is identical to the one uploaded by their frightened victim then yes, it will be blocked. But there is nothing to stop the ex uploading it to another site and linking to it on Facebook, even if it wouldn't appear on Facebook itself.

But what if the ex realises why they have been blocked, and changes the picture slightly? Hashes work for identical pictures. Alex Stamos, chief security officer for Facebook, said that simple changes like re-sizing should not fool the hash. It's not clear whether cropping it, adding a filter or scribbling on the background will create a different hash that will fail to match.

And failure isn't the only problem. What if sharing the images with Facebook actually makes it more likely that they will become public?

The security implications surrounding this are significant. First, just sending an image is a risk. The user is creating a copy of the photo that could be hacked or intercepted, especially if their phone or computer is stolen. Then there is the possibility of human error by the user. It seems likely at least some people will accidentally send the images to someone else in their contact list instead of to themselves.

Finally, there is the enormous issue of how far we can trust Facebook and its staff. This is a small pilot and is likely to be run to tight standards. But further issues will likely appear if it is scaled up. In order to create a hash, the picture has to be seen by a member of Facebook's staff. Antigone Davis, Facebook's global head of safety, says the images will only be seen by "a specially trained representative from our community operations team". Doubtless in the pilot these people will be well vetted.

But if this was to be rolled out to Facebook's billion-strong community of users, this team would have to be enormously expanded. The mind-bending work of ill-paid, ill-trained and replaceable community moderators has been well documented. These people, tempted by the idea of a first-rung on a career at a fashionable company, can be traumatised by the endless viewing of horrific pictures of animal cruelty, car crashes and sexual violence, and often burn out within months.

Can Facebook guarantee that these photos, trustingly uploaded by desperate people trying to break free from damaging relationships, will only be seen by responsible staff? Or will they, over time, be farmed out to subcontractors, trainees and people who are themselves damaged by constant exposure to violence and sex online. However good Facebook's own security is, there would be little to stop a disgruntled, bored or

malicious employee simply taking pictures of their screen with their phone and uploading them to another site.

Then there is the corporation itself. The company has a long history of controversial changes to its terms and conditions, including how they use and retain users' data, even after people have quit the platform. They have fought court cases brought by revenge porn victims who feel failed by the system, in one case by a British 14-year-old.

This is a well-meaning initiative, but it's just not clear that we can trust this commercial organisation to make the right decisions about how they hold this most sensitive data. Stamos has complained that the company gets criticised for imperfect solutions. It's true that partial solutions are better than none, and that pre-emptive solutions are better than clean-ups when the damage is done. But this is a solution that carries its own risk.

This article was originally published on The Conversation. Read the original article.

Provided by The Conversation

Citation: Facebook wants your nude photos to prevent 'revenge porn' – here's why you should be sceptical (2017, November 15) retrieved 2 June 2024 from https://phys.org/news/2017-11-facebook-nude-photos-revenge-porn.html

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.