Joint Committee on Finance, Public Expenditure and Reform and Taoiseach
24 Samhain 2023, 7pm
Opening Statement by Meta

Introduction

Go raibh maith agat a Chathaoirligh. I'd like to thank the members of the Committee for inviting us to today's session to discuss the subject of authorised push payment fraud.

My name is Dualta Ó Broin, I'm Head of Public Policy for Meta in Ireland. I'm joined by my colleague Philip Milton, who is a member of our UK Public Policy team. Philip has been involved in the engagements with the UK Parliament and UK Government on the issue of Fraud for the last few years. In this opening statement, I will focus on Meta's efforts to tackle fraud on our platforms.

The safety of our users is a priority for Meta and we therefore take a zero tolerance approach to fraud on our platforms. By its very nature fraud is adversarial and hard to spot and the perpetrators of fraud are continually searching for ways to subvert the rules, processes and safeguards we put in place to protect our users. The perpetrators are also operating across platforms and industries to avoid disruption by any one platform - which makes this a very challenging area. Just as it is unlikely that fraud will ever be eradicated in society at large, it is unlikely we will ever be able to completely eradicate it online.

Nonetheless Meta is committed to doing all we can to prevent fraudulent activity on our platforms wherever we can. We invest substantially in our safety and security teams to that end. All told since 2016, we've spent about $20bn on teams and technology in this area and that is not slowing down - $5bn of that was in the last year alone.

We have a team of highly trained experts solely focused on identifying fraud and building tools to counter this kind of activity, which are used to help catch suspicious activity at various points of interaction on the site, block accounts used for fraudulent purposes, and remove bad actors.

Policies

It is directly in our interests to do all we can to combat fraud on our platforms. Failure to do so will expose our users to risk, severely degrade the experience of using our platforms for users and make them an unattractive place for brands and businesses to advertise. Meta has a set of strict Advertising Policies, Community Standards and Community Guidelines, which govern what is and is not allowed in advertising and non-paid (organic) surfaces on Facebook and Instagram.

Where we believe anyone has violated our terms, standards, and policies we take action and use a range of tools to enforce our policies, either via proactive automated systems and/or reactive methods.

Our approach to fraud

We deploy a combination of proactive detection and reactive action to disrupt bad actors on our platforms. This includes using our Artificial Intelligence (AI) systems to proactively detect suspicious activity. We focus our attention on behaviours rather than on content, as while the content of these scams changes frequently, the modus operandi of the bad actor typically remains the same.

Where our systems are near-certain that content or profiles are violating (because they possess the signals we associate with a scam to a high degree of confidence) they will immediately be automatically removed. Where less certain, content may be prioritised for our moderation teams to review.

Our aim is to catch bad actors proactively, as early as possible, before they have a chance to engage with users. APP is of course concerned with inauthentic behaviour. When someone looks to create a page or profile we will use our AI to check for signs they are being created by real people and not automated bots. This is because scammers can use bots to help them commit fraud. Accordingly one of the tools we use to combat APP is taking down fake accounts and in H1 of 2023, we removed 1.1bn .

We also use a mixture of nudge behaviour and proactive warnings via messenger to let users know when they are messaging an account which is demonstrating behaviour similar to that we have previously seen from scammers. These accounts have not breached the levels we would need to see to suspend or disable an account but are suspicious enough to warn users about.

For fraudulent activity using advertising on our platforms we also focus on behaviours rather than content, given the ever changing nature of these scams. These efforts are geared towards building more proactive tools to automatically take down this content before it goes live using a combination of AI and human review.

Our systems incorporate signals such as user feedback, fake/compromised account signals and ad content signals and tactics which go into building our proactive detection technology. We've also invested in ensuring our specialised reviewers can understand and identify this content - which by its very nature is hard to spot. Relative to other harms on Facebook, the scams space is more complex and difficult for reviewers to accurately classify, so we have sought to build a more holistic understanding of the abuse over time.

Whilst our aim is to catch content proactively, ideally before users report it to us or interact with it, where users do come across such content we want to make the process of reporting it to us and getting it taken down as easy as possible.

Our in app reporting function is available via the "three dots" that appear in every piece of

posted content and users can report organic content which they consider to be harmful in some way or advertising content which they no longer want to see or think is irrelevant or inappropriate. These reports are an integral part of training our systems to better spot fraudulent activity.

We also have the ability to onboard regulators to our Consumer Policy Channel (CPC). The CPC enables us to work with consumer protection bodies, Government departments, regulators and law enforcement around the world to help us better detect and remove content that violates our policies or local law by taking action on content reported to us by agencies who have the appropriate authority to make determinations in relation to the commercial content or activity they are reporting. We have several of these relationships in Ireland covering a wide range of regulatory issues and harms.

Where we see a trend towards a particular type of activity that is not captured by our policies we review those policies with the input of experts to ensure they remain fit for purpose as the landscape evolves.

Our priority is always to act against a bad actor as quickly as possible for any violation, but we are operating in a particularly adversarial space with bad actors who use increasingly sophisticated means to avoid detection. This is a complex issue that requires a joined up multi-stakeholder approach.

I hope that this provides the members of the committee with an overview of how seriously Meta takes the issue of authorised push payment Fraud and Fraud more generally - and the various methods we employ to combat it. We look forward to the Committee's questions on this important subject.