making communications work
for everyone

Ymosodiad yn Ninas Buffalo: Heriau a Chyfleoedd ar gyfer Diogelwch Ar-lein

Arolwg o rai ymatebion y diwydiant a'r sefyllfa o ran cydweithredu ar draws y diwydiant mewn cysylltiad ag ymosodiadau terfysgol sy'n cael eu ffrydio'n fyw

Dyddiad cyhoeddi: 12 Hydref 2022

Cynnwys

Adran

Crynodeb Gweithredol 1

Crynodeb Gweithredol

Ar 14 Mai 2022, cafwyd ymosodiad yn ninas Buffalo, Efrog Newydd, a arweiniodd at farwolaeth deg o unigolion gan adael tri arall wedi'u hanafu. Cafodd yr ymosodiad ei ffrydio'n fyw ar-lein a chafodd fersiynau o'r fideo eu dosbarthu ar nifer o wasanaethau ar-lein, gan adael defnyddwyr y DU yn agored i gynnwys yn ymwneud â therfysgaeth. Fel rheoleiddiwr llwyfannau rhannu fideos (VSPs) sydd wedi'u lleoli yn y DU yn ogystal â darpar reoleiddiwr Diogelwch Ar-lein y DU, roeddem wedi ceisio dysgu o'r digwyddiad trasig drwy adolygu ymatebion y diwydiant i'r ymosodiadau byw a'r cydweithredu sy'n bodoli ar draws y diwydiant i atal lledaenu cynnwys cysylltiedig.

Mae'r ymosodiad yn Buffalo wedi dangos y gall cynnwys sy'n terfysgol, treisgar a llawn casineb ar-lein chwarae rhan allweddol o ran radicaleiddio defnyddwyr agored i niwed. Yn ôl y sôn, cafodd yr ymosodwr ei ysbrydoli gan yr ymosodiadau ar fosgiau Christchurch yn Seland Newydd, lle collodd 50 o unigolion eu bywydau ac roedd wedi defnyddio'r un dull hefyd. Cafodd rhannau helaeth o faniffesto ymosodwr Christchurch eu copïo, a chyfeiriodd ymosodwr Buffalo ato fwy nag unwaith yn ei ddyddiadur.

Mae ymosodiad o'r fath yn tynnu sylw at y cymhlethdod a'r heriau sy'n gysylltiedig â chynnwys terfysgol ar-lein. Dywedwyd bod yr ymosodwr wedi'i radicaleiddio'n rhannol trwy ddod i gysylltiad â chynnwys hiliol ar y bwrdd negeseuon *4chan*. Mae hefyd yn ymddangos ei fod wedi gweithredu ar ei ben ei hun ac nad oedd yn perthyn i unrhyw sefydliad terfysgol neu sefydliad cysylltiedig arall. Cafodd dolenni i'w ddyddiadur ar-lein eu rhannu mewn gweinyddion preifat a thrwy negeseuon uniongyrchol tua 30 munud cyn yr ymosodiad; cliciodd 15 o unigolion ar y ddolen i'r dyddiadur. Roedd ffrwd fyw'r ymosodiad yn para llai na 2 funud ac roedd 28 o wylwyr, neu lai, wedi gwylio ffrwd fyw'r sianel ar ryw adeg yn ystod y darllediad. Er gwaethaf y nifer isel hwn, cafodd lluniau o'r ymosodiad eu lledaenu ar draws llwyfannau a'u gweld gan filiynau o bobl, a rhannwyd copïau o'r dyddiadur a'r maniffesto yn bennaf trwy lwyfannau llai. Mae sawl agwedd posibl i'r niwed yn sgil defnyddio gwasanaethau ar-lein o'r fath: natur annymunol a graffig y fideo, cymunedau yn gorfod ail-fyw digwyddiadau tebyg i'r rhai a oedd wedi effeithio arnynt yn y gorffennol, a'r risg y bydd defnyddwyr yn agored i gael eu radicaleiddio ar-lein.

Ar ôl dod yn ymwybodol bod yr ymosodiad yn cael ei ffrydio'n fyw ac yn cael ei ledaenu ar lwyfannau rhannu fideos rydym yn eu rheoleiddio ar hyn o bryd, aethom ati i drefnu cyfarfod brys â chynrychiolwyr o'r llwyfannau perthnasol i sefydlu beth yn union oedd wedi digwydd ac a oedd eu systemau a'u prosesau wedi ymateb yn effeithiol i'r ymosodiad a gafodd ei ffrydio'n fyw.

Ymgysylltodd Twitch â ni'n rhagweithiol i rannu gwybodaeth fanwl ac roedd TikTok a BitChute hefyd wedi cydweithredu'n gynhwysfawr â'n hymholiadau. Ar sail asesiadau trylwyr o'r llwyfannau hyn, a'n hystyriaeth i weld a oedd tystiolaeth fod methiannau posibl i gydymffurfio, rydym wedi penderfynu peidio â chynnal unrhyw ymchwiliad gorfodi ffurfiol ar hyn o bryd. Fodd bynnag, byddwn yn parhau i ymgysylltu â'r tri llwyfan rhannu fideos ynghylch y mesurau diogelwch sydd ganddynt ar waith i atal ymosodiadau o'r fath rhag cael eu ffrydio'n fyw ac i ddiogelu defnyddwyr rhag lledaenu cynnwys cysylltiedig.

Cafodd y deunydd a'r cynnwys cysylltiedig eu dosbarthu hefyd ar wasanaethau eraill a oedd yn debygol o ddod o fewn hyd a lled ein gwaith rheoleiddio arfaethedig ar ddiogelwch ar-lein. I'n helpu ni i ddatblygu dealltwriaeth gynhwysfawr o'r ymosodiad, aethom ati i gynnal dros 20 o gyfarfodydd gyda rhanddeiliaid arbenigol gan gynnwys actorion yn y diwydiant, ymchwilwyr, llywodraethau a rheoleiddwyr sydd ag arbenigedd a phrofiad o radicaleiddio a lledaenu cynnwys terfysgol ar-lein.

Rydym wedi nodi pedair her a dewis allweddol sy'n wynebu cwmnïau sy'n ceisio atal terfysgaeth rhag ecsploetio eu gwasanaethau:

- Ers yr ymosodiad yn Christchurch, mae mentrau ar draws y diwydiant fel *Christchurch Call* (CCU), Fforwm Rhyngrwyd Byd-eang ar Wrthderfysgaeth (GIFCT), Fforwm Rhyngrwyd yr UE (EUIF) a *Tech against Terrorism* (TaT) wedi cael effaith nodedig ar y gwaith o leihau nifer y deunydd terfysgol sy'n cael eu ffrydio'n fyw ar-lein, yn enwedig drwy eu protocolau ymateb i argyfwng. Fodd bynnag, nid oes yr un fenter ar draws y diwydiant yn gallu cynnig ateb i'r her fyd-eang hon, yn enwedig os yw eu haelodaeth neu gyfranogiad yn gyfyngedig. Heb gynnwys gwasanaethau ar-lein llai o faint a mwy newydd, efallai na all un model ddarparu ateb cynhwysfawr. Mae'n bosibl bydd angen mwy a mwy o ddulliau gweithredu ar waith ar yr un pryd i fynd i'r afael â gwahanol wasanaethau a gwahanol fathau o fygythiadau.

- Mae llwyfannau'n defnyddio dulliau gwahanol o adnabod actorion terfysgol, eithafol treisgar neu gas. Nid yw rhai o'r llwyfannau hyn ond yn gwahardd sefydliadau sydd wedi'u gwahardd gan wladwriaethau neu sefydliadau rhynglywodraethol. Mae eraill yn defnyddio dull gweithredu mwy eang, gan wahardd amrywiaeth ehangach o grwpiau a chynnwys, gan gynnwys iaith casineb, deunydd sy'n cymell trais, a thwyllwybodaeth/theorïau sy'n ceisio radicaleiddio defnyddwyr. Gall gwasanaethau osod eu telerau ac amodau eu hunain ar yr amod eu bod yn bodloni unrhyw amodau penodol sy'n ofynnol gan ddeddfau a rheoliadau perthnasol. Ond wrth wneud hynny, mae'n anochel eu bod yn gwneud cyfaddawd rhwng diogelu eu defnyddwyr rhag cynnwys a allai fod yn niweidiol, a'u gallu i ddweud eu dweud. Er mwyn i ddefnyddwyr wybod yr hyn y gallant neu na allant ei rannu neu ddod ar ei draws ar lwyfannau, mae angen i'r llwyfannau sicrhau bod eu telerau gwasanaeth yn hygyrch ac yn egluro sut maent yn diffinio ac yn mynd i'r afael â chynnwys terfysgol, treisgar a chas.

- Mae offer riportio defnyddwyr yn dal i chwarae rhan bwysig yn y gwaith o dynnu sylw at gynnwys terfysgol, treisgar a chas mewn ffrydiau byw, ac mae defnyddwyr yn fwyaf tebygol o'u defnyddio pan maent yn fwy hygyrch ac yn haws eu defnyddio. Byddai cefnogi hyn gyda thimau cymedroli sydd â'r adnoddau priodol hefyd yn cryfhau ei effeithlonrwydd.

- Er mai prin yw ymosodiadau terfysgol sy'n cael eu ffrydio'n fyw, maent yn ddigwyddiadau annymunol dros ben ac yn peri risgiau sylweddol i ddefnyddwyr. Mae cyfleoedd ar gael i lwyfannau leihau'r risg o'r digwyddiadau hyn, a hynny drwy gyflwyno nodweddion sy'n cyfyngu ar fynediad at ffrydio byw mewn amgylchiadau penodol. Rydym yn annog llwyfannau i archwilio'r risgiau a achosir gan eu gwasanaethau ac i wneud ymdrechion priodol i'w gwneud eu hunain yn gadarn er mwyn osgoi cael eu hecsbloetio gan ymosodwyr.

Mae'r gwaith hwn wedi helpu i ddatblygu ein dealltwriaeth o'r ffyrdd y mae gweithredwyr terfysgol ac eithafol treisgar yn defnyddio ffrydio byw i ledaenu deunydd o'u hymosodiadau, yn ogystal ag o effaith ymatebion a chamau cydweithredol y diwydiant. Byddwn yn ystyried y canfyddiadau hyn fel y bo'n briodol yn ein gwaith parhaus tuag at lansio rheoliadau Diogelwch Ar-lein y DU, sy'n cynnwys paratoi ein hasesiad risg sector, canllawiau asesu risg ar gyfer gwasanaethau rheoleiddiedig a chodau ymarfer sy'n ymwneud â chynnwys anghyfreithlon. Rydym yn bwriadu eu cyhoeddi ar ffurf drafft yng ngwanwyn 2023 ar gyfer ymgynghori arnynt, ar ôl i'r Bil Diogelwch Ar-lein gael ei basio. Rhwng nawr a hynny, byddwn yn parhau i weithio gyda VSPs i sicrhau bod ganddynt fesurau priodol i ddiogelu defnyddwyr rhag deunydd terfysgol anghyfreithlon a chynnwys sy'n cymell casineb a thrais. Rydym hefyd yn disgwyl cyhoeddi rhagor o ymchwil i'r risgiau y mae terfysgaeth, trais a chasineb yn ei achosi i ddefnyddwyr ar-lein y DU yn ystod y flwyddyn nesaf.