

Guia docent amb activitats a l'aula pel Dia de la Internet Segura

DIS2024



OBJECTIU DE LA GUIA

La guia docent sobre la tecnologia d'hipertrucatge o *deep fake* amb activitats a l'aula, elaborada en el marc del Dia de la Internet Segura 2024, pretén ser **una eina de suport al professorat** per donar a conèixer i sensibilitzar l'alumnat sobre aquesta tecnologia realitzada amb intel·ligència artificial (IA) i que presenta riscos en l'àmbit de la ciberseguretat.

Es preveu que en els pròxims mesos hi hagi una allau de *deep fakes* de vídeo i d'àudio a les xarxes socials. Segons el Fòrum Econòmic Mundial, la taxa de creixement anual d'aquest tipus de contingut audiovisual és del 900%. Mentre les xarxes s'inunden d'aquest tipus de contingut, la realitat és que **el 71% de la població mundial no sap què és un *deep fake*** o hipertrucatge. En concret, a l'estat espanyol aquesta xifra s'enfila fins al 75%.

En relació al tipus de contingut *deep fake* que s'està creant **més del 95% està relacionat amb pornografia no consensuada**. Es tracta de material pornogràfic en què els creadors agafen cares que troben a Internet, a les xarxes socials, a vegades d'infants, i les fan servir per crear imatges pornogràfiques o pedòfiles falses.

En aquest context, **la conscienciació i l'educació són claus per identificar i combatre aquest fenomen** emergent. Per això, és necessari inculcar als infants i joves un esperit crític i donar-los eines perquè aprenguin a distingir entre allò que és fals i el què no, a més de sensibilitzar-los sobre el risc de fer servir aquest tipus de tecnologia.

QUÈ ÉS UN DEEP FAKE O HIPERTRUCATGE?

L'aparició de la intel·ligència artificial generativa ha provocat que la producció de continguts *deep fakes* sigui fàcil i barata, pràcticament a l'abast de qualsevol usuari. Però què és exactament un hipertrucatge? Es tracta d'una **tècnica de manipulació audiovisual que fa servir algoritmes d'intel·ligència artificial** per crear muntatges molt realistes que contenen situacions fictícies. L'objectiu d'aquest contingut és sovint fer difusió de notícies falses, burla de polítics o persones rellevants, crear continguts humorístics, fer material pornogràfic, etc. Però no hem d'oblidar que també hi ha hipertrucatges positius, és a dir, gent que fa servir la tecnologia per fer accions de sensibilització o divulgació de continguts culturals, socials o didàctics, entre d'altres.

Els hipertrucatges poden ser de veu, d'imatge o de vídeo i una de les tècniques més usades és la **superposició de diferents identitats** que permet fer muntatges molt realistes amb persones dient coses o intervenint en escenes on no han estat mai.

Cal destacar també el concepte d'**humans sintètics**, éssers dissenyats per ordinador que pretenen semblar-se i comportar-se com a humans reals i que es poden fer servir per molts propòsits: en educació o investigació, però també, per exemple, en l'entreteniment.

COM PODEM DETECTAR UN HIPERTRUCATGE?

Tot i que sovint els *deep fakes* poden semblar evidents, altres vegades no ho són tant. És aquí on cal observar amb atenció les imatges i detectar errors o incongruències.

- A les **imatges** ens hem de fixar bé en les cares: molts hipertrucatges són transformacions facials. Per això hem de fixar-nos en els ulls i la pell i en especial les galtes i el front. Hi ha altres elements que ens poden donar pistes, com ara els reflexes de les ulleres o la barba.
- Als **vídeos** podem detectar si hi ha hipertrucatge si ens fixem en el parpelleig i el moviment dels llavis. També és important analitzar si les imatges estan pixelades i captar si hi ha alteracions en les cares i el fons de les imatges. La durada del vídeo també ens pot ajudar a interpretar si és real o fals: la majoria de *deep fakes* són vídeos molts curts.
- En un **àudio**, els *deep fakes* es poden detectar si hi ha pronunciació incorrecta, sons metàl·lics o monòtons o sons pocs naturals. L'hipertrucatge es fa més difícil de reconèixer si l'àudio va acompanyat d'efectes sonors o si es tracta d'una cançó.

MALS USOS DE L'HIPERTRUCATGE

La **desinformació** és un dels mals usos que es pot fer d'aquesta tecnologia. En aquesta imatge feta amb intel·ligència artificial (IA), un periodista nord-americà va decidir recrear una escena falsa de detenció de l'expresident dels Estats Units Donald Trump. Tot i que la intenció del creador de la imatge era que fos un simple joc, la imatge es va fer viral, va fer la volta al món i moltes persones van pensar que realment havia estat detingut.

La **sextorsió** o nus falsos és un altre exemple de mala pràctica de la tecnologia deep fake. Existeixen diverses aplicacions que fan servir la IA i xarxes neuronals per eliminar la roba i recrear nus molt realistes.

També hi ha ciberdelinqüents que aprofiten aquesta tecnologia per cometre **estafes amb hipertrucatge de veu** o generar contingut fals. Un usuari de TikTok va recrear la veu del cantant de reggaeton Duki posant la seva veu a temes d'altres artistes. Les recreacions, tot i ser falses i vulnerar els drets d'autor, acumulen milions de reproduccions.



BONS USOS DE L'HIPERTRUCATGE

La tecnologia deep fake, com qualsevol eina, no és negativa ni perillosa per si sola: hi ha casos d'hipertrucatges que s'han fet servir per **sensibilitzar la població**. Per exemple, en aquest cas, la companyia telefònica francesa Orange va publicar un vídeo on es veia els jugadors masculins de la selecció francesa mostrant el seu talent futbolístic. Al final del vídeo, es mostrava que les imatges eren falses: era un hipertrucatge i les imatges originals eren de jugadores de la selecció francesa de futbol. D'aquesta forma, es pretenia conscienciar sobre la desigualtat de gènere en aquest esport.

La **recreació de personatges històrics** és un altre cas d'hipertrucatge positiu. El Museu Salvador Dalí de Florida, als Estats Units, va elaborar un vídeo deep fake on 'reviu' el pintor amb la intenció de divulgar la seva obra. Dalí fa d'amfitrió, saluda els visitants i fins i tot es fa una autofoto o *selfie* amb ells.

L'hipertrucatge també es pot fer servir per fer **divulgació** o explicar conceptes que poden ser difícils d'entendre. Per exemple, Campofrío va fer un anunci fent servir elements *deep fake* per recrear un dinar de Nadal on es posa en context l'hipertrucatge: s'expliquen les limitacions i deficiències de la IA, però també es posen sobre la taula les seves virtuts i les oportunitats que pot oferir.





ACTIVITATS A L'AULA PER AL DIS2024

Per als cicles de **Primària** s'ha elaborat una pàgina interactiva [\[enllaç\]](#) que consisteix en què els alumnes vegin imatges generades per intel·ligència artificial i puguin reflexionar sobre la tecnologia d'hipertrucatge. L'activitat té la missió que els infants puguin **prendre consciència de l'existència de les deep fakes**, ajudar-los a contextualitzar les imatges per detectar hipertrucatges senzills i obrir un debat.

A l'educació **Secundària** es proposa que utilitzeu la pàgina interactiva [\[enllaç\]](#) per **sensibilitzar els adolescents sobre els tipus de continguts** que es poden generar a través de la intel·ligència artificial, així com donar-los unes nocions bàsiques per a detectar l'hipertrucatge i explicar els riscos associats a l'ús d'aquesta tecnologia.

ALTRES PROPOSTES PER TRASLLADAR EL MATERIAL A L'AULA

Per tal de conscienciar i sensibilitzar sobre l'hipertrucatge i poder fer activitats relacionades, es poden fer diverses activitats adaptades a cada tipus d'hipertrucatge:

- **Deep fake d'imatge:** Es pot plantejar a l'alumnat que existeixen imatges de persones no reals, és a dir, sintètiques. Per exemple [a través d'aquest web](#) es generen cares falses. Aquest tipus de material també es pot utilitzar per il·lustrar fets que no han passat, per exemple les fotografies sobre la [detenció de Donald Trump](#) o aquesta [imatge del Papa Francesc](#).
- **Deep fake de vídeo:** Per tractar aquest tema es poden proposar exemples positius amb anuncis, com per exemple el del mundial femení de futbol o [personatges de pel·lícules](#). També es pot parlar sobre els [usos als videojocs](#), construint personatges amb intel·ligència artificial que donen suport al jugador. Després d'analitzar l'hipertrucatge de vídeo es pot enllaçar amb els mals usos d'aquesta tecnologia: per exemple [l'extorsió pornogràfica](#) o sextorsió.
- **Deep fake d'àudio:** Els alumnes poden escoltar [la cançó deep fake de Duki](#) o, en el cas de Primària, el muntatge de Freddie Mercury cantant [Let It Go](#). A partir d'aquí es proposen diverses reflexions, sobre si els agrada més o menys i posteriorment introduir l'element que aquesta tecnologia pot suposar un frau i una violació dels drets d'autor.



MÉS INFORMACIÓ

A les nostres xarxes socials, **Facebook** , **X** i **Instagram**, i a la web d'**Internet Segura** trobareu més contingut en els pròxims dies. També us invitem a seguir i utilitzar els **hashtags**: #TrobaEIFake #DiaInternetSegura #SID2024 #SaferInternetDay.

Junts, podem difondre el missatge més àmpliament!

CATALUNYA
ENS CONNECTA

