Чим загрожують глибокі фейки і як їх розпізнати?

Протягом трьох років успішний журналіст тижневика «The New Republic» (https://newrepublic.com/) Стівен Ґлас сфабрикував щонайменше 27 текстів. Його аферу викрили після статті «Hack Heaven» у травні 1998 року. Під час її написання він створив сайт неіснуючої фірми, підробив повідомлення в голосовій пошті і навіть залишив записи в своєму блокноті, аби створити ілюзію справжньої журналістської праці.

Викрити підробки вдалося журналістові Forbes, який зайнявся фактчекінгом аби дізнатися джерела сенсаційного матеріалу. Всі ці події відображені в біографічному фільмі «Афера Стівена Ґласа» (Shattered Glass). Після цього випадку кожного репортера зобов'язали підтверджувати свої тексти фотографіями. Мовляв, це унеможливить нечесність.

Але чи це так? Письменниця Сьюзан Зонтаг вважає, що фотографії — ненадійні оповідачі. «Незважаючи на презумпцію правдивості, яка додає всім світлинам авторитету, праця, яку роблять фотографи, не є загальним винятком із сумнівної торгівлі між мистецтвом і правдою», — пише Сьюзан http://bit.ly/357GR6c

І вона частково права. Нині підробити зображення може ледве не кожен, у кого є фотошоп або інший фоторедактор. Не встигли ми навчитися боротися з фотофейками, зробленими за допомогою фотошопу, як з'явилася нова, більш масштабна й вражаюча технологія. Якщо раніше правдоподібно сфальсифікувати відео могли хіба що голівудські майстри зі спецефектів із багатомільйонними бюджетами, то тепер це стає під силу навіть людям без особливих навичок і дорогого обладнання.

І в такому здешевленні ховається найбільша загроза. Кожен за бажання може створити подію, якої ніколи не було. Її можна просто згенерувати за допомогою спеціальних комп'ютерних

алгоритмів, нейронних мереж і штучного інтелекту. Таке явище називають Deepfake – глибокою підробкою, лицедурством. Це відеоролики, на яких до тіла однієї людини «причепили» обличчя іншої, і зробили це таким чином, що відрізнити підробку неймовірно складно.



Межа між дійсністю і вигадкою стає дедалі примарнішою. Своїм очам і вухам уже не можна беззастережно довіряти. Перегляньте ці
відео http://bit.ly/2OrtQOM, http://bit.ly/2LSDyYG, http://bit.ly/2InNLKI Щоб таке спростувати потрібно мати неабиякий хист. Як його набути, раніше ми вже докладно розповідали http://bit.ly/2It73yr

Знавці прогнозують, що цей високотехнологічний спосіб «вкладати слова в чиїсь уста» стане найсучаснішою зброєю в

умовах інформаційних воєн http://bit.ly/31Lu5rK
У майбутньому «глибокі фейки» можуть мати колосальні наслідки — від остаточної втрати суспільної довіри до медіа до впливу на результати народного волевиявлення. Простіше кажучи, люди будуть розгублені, збиті з пантелику. Діятиме презумпція брехні — потрібно буде не спростовувати фейки, а доводити, що написане/показане — правда.

У деяких країнах уже всерйоз занепокоїлися цією проблемою. Наприклад, нещодавно в Каліфорнії ухвалили два закони, які обмежують використання технології Deepfake https://cnet.co/2oZftqe Перший документ забороняє розповсюджувати маніпулятивні відеоролики з кандидатами протягом двох місяців перед виборами. Інший дозволяє мешканцям штату подавати до суду, якщо хтось використав їхній образ у відео-deepfake порнографічного характеру.

Facebook, Microsoft і багато інших компаній пропонують винагороду тим, хто розробить технологію, яка зможе розпізнавати відео, створені за допомогою нейромереж, і запобігати їхньому поширенню. Для цього був створений спеціальний фонд Deepfake Detection Challenge, в якому наразі вже є $10 млн. Конкурс на отримання відповідного гранту стартує в грудні цього року https://tcrn.ch/2AO8Qd2

Хай там як, із розвитком технологій вчимося і ми. Згадайте перші частини «Термінатора» чи «Зоряних воєн». Колись вони здавалося настільки майстерними, що хотілося вірити кожному кадру. Сьогодні ж усе це сприймається як «вирвиоко» — помітно іграшковість роботів, недолугість лазерів, халтурність інших спецефектів.

Можливо за якихось десять років нам не знадобляться спеціальні програми для розпізнавання нинішніх підробок. Достатньо буде побіжного погляду. Як от, наприклад, зі світлинами фей, про які ми недавно писали http://bit.ly/2AVHdyB Свого часу їх визнали найбільшою містифікацією XX століття. Сьогодні ж навіть не віриться, що колись люди могли у таке вірити.