ΠΗΓΗ: THE EUROPEAN LIBERAL FORUM
https://liberalforum.eu/event/on-the-agenda-how-can-the-eu-build-a-future-proof-legal-framework-to-tackle-deepfakes/

ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ
ΔΠΜΣ ΔΙΚΑΙΟ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗ

DEEPFAKES: ΤΕΧΝΟΛΟΓΙΚΗ ΚΑΙ ΝΟΜΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

ΠΑΝΑΓΙΩΤΑ ΠΑΝΑΓΟΠΟΥΛΟΥ
ΜΑΡΤΙΟΣ 2023

Καθ. ΘΕΟΧΑΡΗΣ ΔΑΛΑΚΟΥΡΑΣ
Καθ. ΚΩΝΣΤΑΝΤΙΝΟΣ ΨΑΝΝΗΣ

ΕΙΣΑΓΩΓΗ

ΑΝΤΙΚΕΙΜΕΝΟ ΕΡΓΑΣΙΑΣ
Η ΤΕΧΝΟΛΟΓΙΑ DEEPFAKE

ΔΟΜΗ ΕΡΓΑΣΙΑΣ
Α' ΜΕΡΟΣ: ΤΕΧΝΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ
Β' ΜΕΡΟΣ: ΝΟΜΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ
DEEP + FAKE = DEEP LEARNING + ΨΕΥΤΙΚΟ - ΠΛΑΣΤΟ

ΣΚΟΠΟΣ
ΤΕΧΝΟΛΟΓΙΕΣ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΟΥΝΤΑΙ – ΠΩΣ ΔΗΜΙΟΥΡΓΕΙΤΑΙ ΕΝΑ DEEPFAKE

ΤΕΧΝΗΤΑ ΜΕΣΑ ΟΠΟΥ ΕΝΑΣ ΑΛΓΟΡΙΘΜΟΣ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ «ΜΑΘΑΙΝΕΙ» ΕΝΑ ΠΡΟΣΩΠΟ ΚΑΙ ΤΟ ΑΝΤΙΚΑΘΙΣΤΑ ΜΕ ΑΛΛΟ

ΣΚΟΠΟΣ
ΣΚΟΤΕΙΝΗ ΠΤΥΧΗ ΤΗΣ ΤΕΧΝΟΛΟΓΙΑΣ – ΖΗΤΗΜΑΤΑ ΤΕΛΕΣΗΣ ΠΟΙΝΙΚΩΝ ΑΔΙΚΗΜΑΤΩΝ

ΤΟ ΠΡΟΒΛΗΜΑ
Η ΤΕΧΝΟΛΟΓΙΑ ΩΣ ΟΠΛΟ ΣΤΑ ΧΕΡΙΑ ΕΝΟΣ ΚΑΚΟΒΟΥΛΟΥ ΧΡΗΣΤΗ

ΜΕΘΟΔΟΛΟΓΙΑ
ΠΡΟΣΦΑΤΕΣ ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ (2018-2022)

ΣΗΜΑΝΤΙΚΟΤΗΤΑ ΤΟΥ ΘΕΜΑΤΟΣ
ΚΟΙΝΩΝΙΚΟΗΘΙΚΑ ΚΑΙ ΝΟΜΙΚΑ ΖΗΤΗΜΑΤΑ

ΣΥΜΠΕΡΑΣΜΑ – ΠΡΟΤΑΣΕΙΣ
ΟΦΕΛΟΣ Ή ΑΠΕΙΛΗ;
ΠΡΟΤΑΣΕΙΣ ΑΝΤΙΜΕΤΩΠΙΣΗΣ

ΕΙΣΑΓΩΓΗ ΣΤΗΝ ΤΕΧΝΟΛΟΓΙΑ ΤΩΝ DEEPFAKES
ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ

ΠΗΓΗ: National Institute of Standards and Technology https://www.nist.gov/internet-things-iot
@Blue Planet Studio

- Τεχνητή νοημοσύνη μια πρακτική πραγματικότητα όχι επιστημονική φαντασία.
- Διαδίκτυο των Πραγμάτων (Internet of Things - IoT) όλα τα καθημερινά "πράγματα" συνδέονται με το διαδίκτυο και επικοινωνούν μεταξύ τους ανταλλάσσοντας πληροφορίες και δεδομένα.
- Μεγάλος όγκος δεδομένων από πολλές συσκευές
- Ανησυχία για την ασφάλεια των δεδομένων

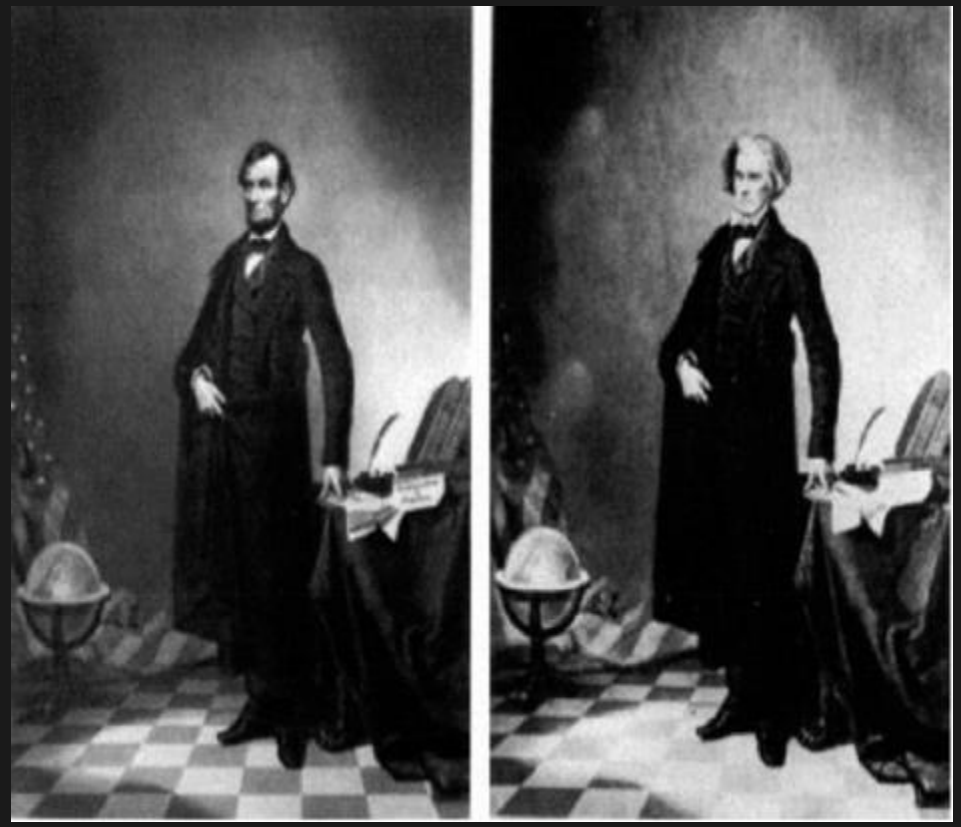
- Εντυπωσιακή ανάπτυξη των μέσων κοινωνικής δικτύωσης ως μέσο διάδοσης ειδήσεων.
- Το βίντεο και η εικόνα το κυριότερο και πιο δυναμικό μέσο επικοινωνίας.
- Τεράστιες επιπτώσεις στο πως παράγουμε περιεχόμενο, επικοινωνούμε και αντιλαμβανόμαστε τον κόσμο.
- Συνθετικά μέσα (AI) → ποιότητα, εύκολη πρόσβαση σε λογισμικό δημιουργίας, δωρεάν και φτηνή παραγωγή.
- Δισεκατομμύρια χρήστες κινητών τηλεφώνων → δεν βλέπουν μόνο online videos, αλλά δημιουργούν και κοινοποιούν.

ΠΗΓΗ: Η αποχή από τα social media μάς φτιάχνει τη διάθεση – Αποκαλυπτική μελέτη https://www.in.gr/2022/05/06/health/body/apoxi-apo-ta-social-media-mas-ftiaxnei-ti-diathesi-apokalyptiki-meleti/

ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ

Εμφάνιση της φωτογραφίας το 19ο αιώνα → μέσο χειραγώγησης από τον άνθρωπο

ΠΗΓΗ: uploaded by Savita Walia on Researchgate https://www.researchgate.net/figure/Lincolns-head-over-Southern-politician-John-Calhouns-body_fig1_325665869



Το 1865, όταν δολοφονήθηκε ο Αβραάμ Λίνκολν, δεν υπήρχαν εικόνες του Προέδρου σε "ηρωικό στυλ" και για να αντιμετωπιστεί αυτό, ένας χαράκτης αποφάσισε σε μια φωτογραφία να τοποθετήσει το κεφάλι του Λίνκολν πάνω στο σώμα του πολιτικού John C Calhoun. Για έναν αιώνα, κανείς δεν το πρόσεξε. Μόλις πρόσφατα αποκαλύφθηκε ότι η φωτογραφία ήταν παραποιημένη.

ΠΗΓΗ: Schick N. Deepfakes : The Coming Infocalypse. First U.S. ed. New York: Twelve; 2020, σελ. 28

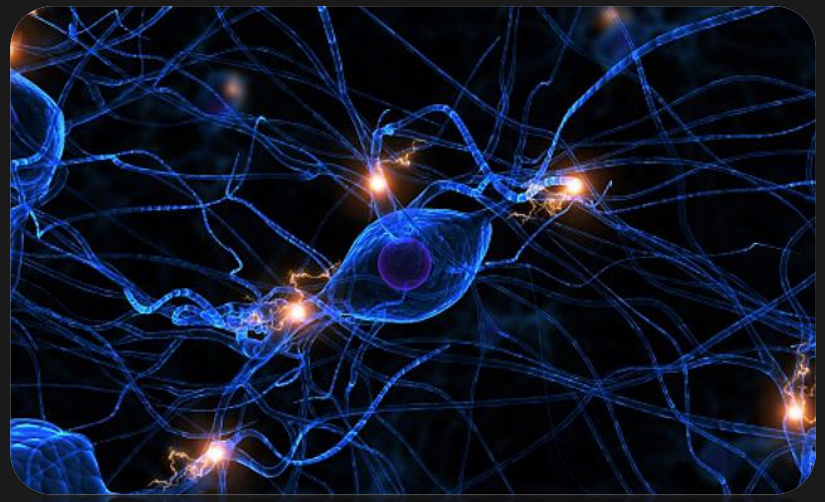
Πάνω: Συνδιάσκεψη Κόμματος Απρίλιος 1925

Κάτω: Αναπαραγωγή φωτογραφίας 1939

Στην πορεία, τη δεκαετία του 1930, μεταξύ των όσων συνέβησαν στο όνομα του Ιωσήφ Στάλιν και της ιδεολογίας του, αναπτύχθηκε μια ολόκληρη βιομηχανία αφιερωμένη στην επεξεργασία φωτογραφιών. Πολλοί πολιτικοί εχθροί του Στάλιν δολοφονήθηκαν ή φυλακίστηκαν και ταυτόχρονα ως δια μαγείας αφαιρούνταν από τις φωτογραφίες.

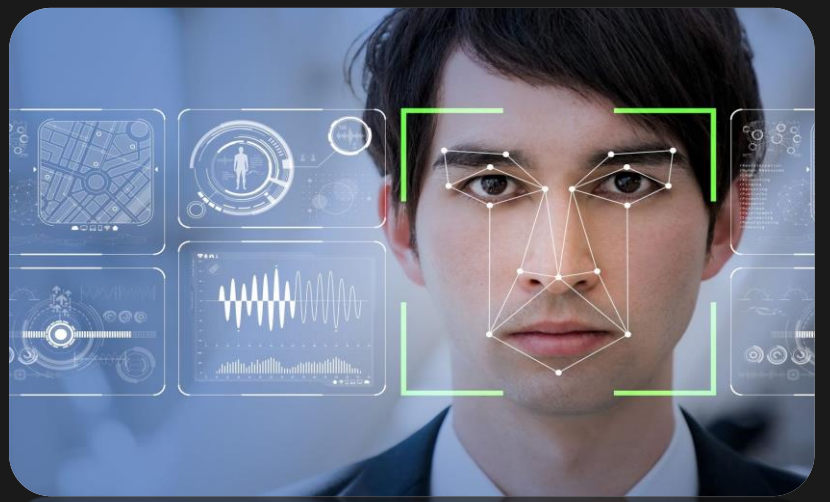
ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ

- Εφεύρεση υπολογιστών 1950 → χαρακτηριστικά νοημοσύνης - ή τεχνητής νοημοσύνης (AI) μιμούμενοι τα νευρωνικά δίκτυα του ανθρώπινου εγκεφάλου.



ΠΗΓΗ: https://www.cea.gr/%CF%84%CE%B5%CF%87%CE%BD%CE%B7%CF%84%CE%AC-%CE%BD%CE%B5%CF%85%CF%81%CF%89%CE%BD%CE%B9%CE%BA%CE%AC-%CE%B4%CE%AF%CE%BA%CF%84%CF%85%CE%B1-%CF%84%CE%BF-%CE%BC%CE%AD%CE%BB%CE%BB%CE%BF%CE%BD-%CF%84%CE%B7/

- Ανταλλαγή προσώπων με πειστικά αποτελέσματα εξαιτίας των τεχνολογιών που χρησιμοποιούνται.
- Δύναμη της Τεχνητής Νοημοσύνης να κάνει τους ανθρώπους να λένε και να κάνουν πράγματα που ποτέ δεν είπαν και δεν έκαναν.
- Κάπως έτσι γεννήθηκε και το πρώτο deepfake. Η ΤΝ και τα deepfake ίσως να είναι η πιο πρόσφατη υπαρκτή απειλή.



ΠΗΓΗ: http://www.securnet.gr/2020/06/blog-post_94.html

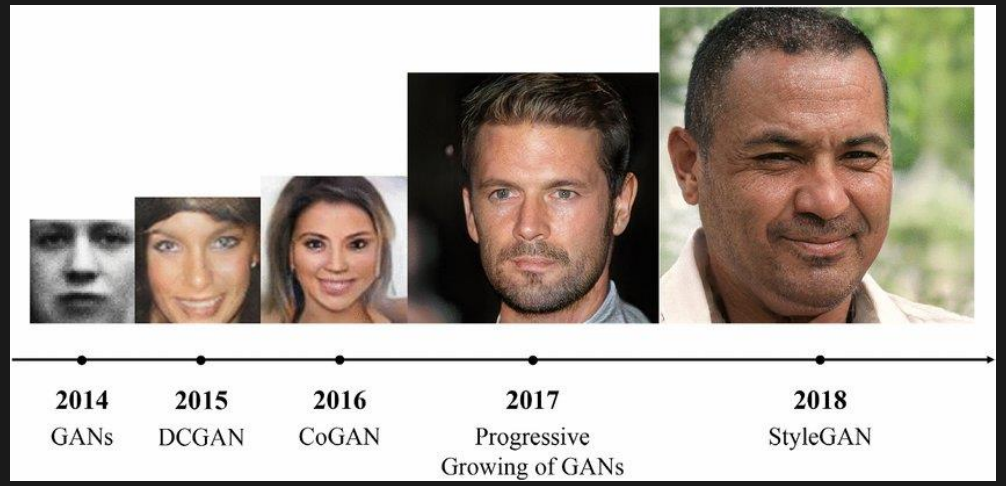
- Το πρόσωπο είναι το πιο χαρακτηριστικό γνώρισμα του ανθρώπου.
- Τεράστια ανάπτυξη της τεχνολογίας αναγνώρισης προσώπου και της σύνθεσης προσώπων.

ΠΗΓΗ: https://www.maxmag.gr/science/anagnorisi-prosopoy-asfaleia-i-apeili-tis-idiotikotitas/

ΤΑ ΓΕΝΕΤΙΚΑ ΑΝΤΙΠΑΡΑΘΕΤΙΚΑ ΔΙΚΤΥΑ ΤΟΥ IAN GOODFELLOW

- ✓ Τεχνητή Νοημοσύνη → Τεχνητά Νευρωνικά Δίκτυα → Βαθιά Μάθηση → Αυτόματοι Κωδικοποιητές
- ✓ Το 2014 ο Αμερικανός ερευνητής Ian J. Goodfellow εφηύρε τα **GAN – Γεννητικά αντιπαραθετικά δίκτυα**.
- ✓ Προγραμμάτισε **δύο δίκτυα βαθιάς μάθησης μαζί σε ένα αντίπαλο παιχνίδι,** για να δημιουργήσει ανθρώπινα πρόσωπα.
- ✓ Το ένα η **γεννήτρια** θα προσπαθούσε να δημιουργήσει νέες πληροφορίες και το άλλο ο **ανιχνευτής** θα προσπαθούσε να τις ανιχνεύσει σε μια συνεχή επαναληπτική διαδικασία.
- ✓ Μέσα σε λίγες ώρες, το σύστημα που είχε δημιουργήσει παρήγαγε ανθρώπινα πρόσωπα που ήταν καλύτερα από οτιδήποτε άλλο είχε φτιάξει η ΤΝ πριν.

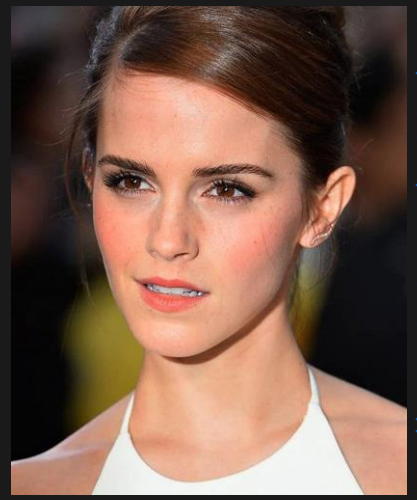
ΠΗΓΗ: https://twitter.com/goodfellow_ian/status/1084973596236144640



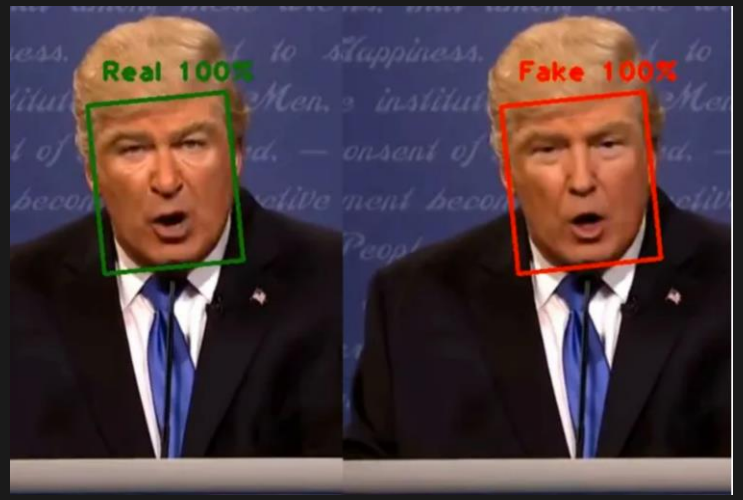
Η ΔΗΜΙΟΥΡΓΙΑ ΤΩΝ ΠΡΩΤΩΝ DEEPFAKES

- ✓ Το φαινόμενο απέκτησε το όνομά του από έναν χρήστη της πλατφόρμας **Reddit**, ο οποίος χρησιμοποιούσε το όνομα **"deepfakes"** (deep learning + fakes) στις **2 Νοεμβρίου 2017**.
- ✓ Ο χρήστης μοιράστηκε τα πρώτα deepfakes τοποθετώντας τα πρόσωπα διασημοτήτων σε βίντεο πορνογραφικού περιεχομένου κυρίως. Αυτό προκάλεσε ευρύ ενδιαφέρον στην κοινότητα του Reddit και οδήγησε σε μια **έκρηξη ψεύτικου περιεχομένου.**
- ✓ Οι πρώτοι στόχοι των deepfakes ήταν διάσημοι άνθρωποι, συμπεριλαμβανομένων ηθοποιών (π.χ. Emma Watson και Scarlett Johansson), τραγουδιστών (π.χ. Katy Perry) και πολιτικών (π.χ. Πρόεδρος Obama, Donald Trump, Volodimir Zelenski).

https://www.imdb.com/name/nm0914612/



https://nationaltoday.com/birthday/scarlett-johansson/

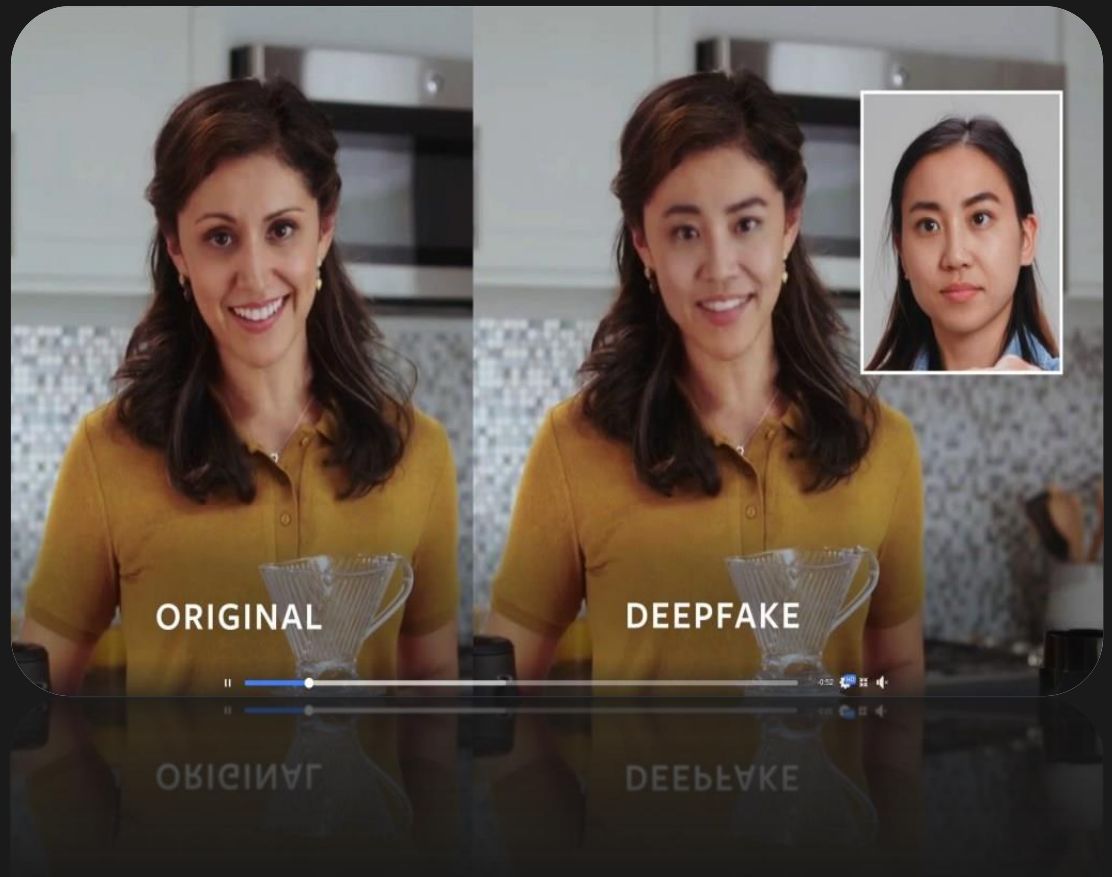


https://spectrum.ieee.org/will-deepfakes-detection-be-ready-for-2020

https://www.bitdefender.com/blog/hotforsecurity/deepfake-president-zelensky-calls-on-ukraine-to-surrender-as-tv-station-hacked/

ΟΡΙΣΜΟΣ ΤΩΝ DEEPFAKES

- ✓ Ένας αναδυόμενος υποτομέας της ΤΝ στην οποία το πρόσωπο ενός ατόμου επικαλύπτεται πάνω από το πρόσωπο ενός άλλου ατόμου, και μέσω των πολλαπλών μεθόδων που βασίζονται στα γεννητικά αντιπαραθετικά δίκτυα (GANs) μπορούν να παραχθούν εικόνες υψηλής ανάλυσης.
- ✓ Μια τεχνολογία παραγωγής πλαστού οπτικοακουστικού υλικού, η οποία με τη βοήθεια της ΤΝ αντικαθιστά το πρόσωπο ενός ατόμου με αυτό ενός άλλου (ανταλλαγή προσώπων – superimposing), με απόλυτη ακρίβεια στο συγχρονισμό χειλιών (lip-sync), στις εκφράσεις του προσώπου και στις κινήσεις των ματιών και του κεφαλιού (puppetmaster).
- ✓ Πλαστό περιεχόμενο, που στόχο έχει να πείσει ότι είναι μια αληθινή παρουσίαση, βάζοντας τους ανθρώπους να λένε και να κάνουν πράγματα που δεν είπαν και δεν έκαναν ποτέ.
- ✓ Δεν χρειάζεται η καθοδήγηση ενός έμπειρου επαγγελματία. Ακόμη και ένας ερασιτέχνης μπορεί να κάνει το ψεύτικο να μοιάζει με αληθινό.



ΠΗΓΗ: https://spectrum.ieee.org/facebook-ai-launches-its-deepfake-detection-challenge

ΤΕΧΝΟΛΟΓΙΕΣ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΟΥΝΤΑΙ

ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

Τομέας της επιστήμης των υπολογιστών, που ασχολείται με τη σχεδίαση ευφυών (νοημόνων) υπολογιστικών συστημάτων, δηλαδή συστημάτων που επιδεικνύουν χαρακτηριστικά που σχετίζονται με τη νοημοσύνη στην ανθρώπινη συμπεριφορά.

ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Ένα υποσύνολο της τεχνητής νοημοσύνης κατά την οποία ένα υπολογιστικό σύστημα δημιουργεί από ένα σύνολο δεδομένων μοντέλα και πρότυπα ώστε οι μηχανές να βελτιώνονται σε εργασίες με την εμπειρία.

ΒΑΘΙΑ ΜΑΘΗΣΗ

Το υποσύνολο της μηχανικής μάθησης που αποτελείται από αλγορίθμους που επιτρέπουν στο λογισμικό να εκπαιδεύεται να εκτελεί εργασίες, όπως η αναγνώριση ομιλίας και εικόνας, εκθέτοντας πολυεπίπεδα νευρωνικά δίκτυα σε τεράστιες ποσότητες δεδομένων.

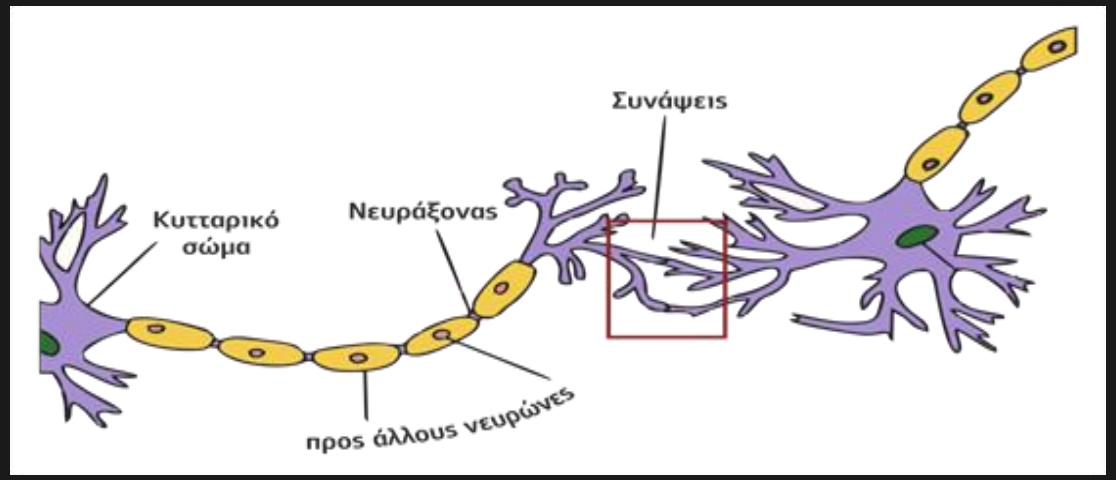
ΠΗΓΗ ΦΟΝΤΟΥ: https://www.kathimerini.gr/world/561906817/google-techniti-noimosyni-ginetai-ena-programma-na-apoktisei-syneidisi-kai-prosopikotita/

ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

ΒΙΟΛΟΓΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

- ✓ **Σώμα** (body) - ο πυρήνας
- ✓ **Δενδρίτες** (dendrites), σημεία εισόδου - μέσο λήψης σημάτων από γειτονικούς νευρώνες
- ✓ **Άξονας** (axon) - μέσο σύνδεσης - σημείο εξόδου
- ✓ **Σύναψη** (synapse) σε κάθε δενδρίτη σαν κενό, η οποία μέσω χημικών αντιδράσεων μεταβάλλει την αγωγιμότητα του νευρώνα, επιταχύνοντας ή επιβραδύνοντας τη ροή ηλεκτρικών φορτίων προς το σώμα του νευρώνα
- ✓ Μέσω των δενδριτών τα ηλεκτρικά σήματα συνδυάζονται και το αποτέλεσμα του συνδυασμού αυτού διαδίδεται μέσω του άξονα προς άλλους νευρώνες.

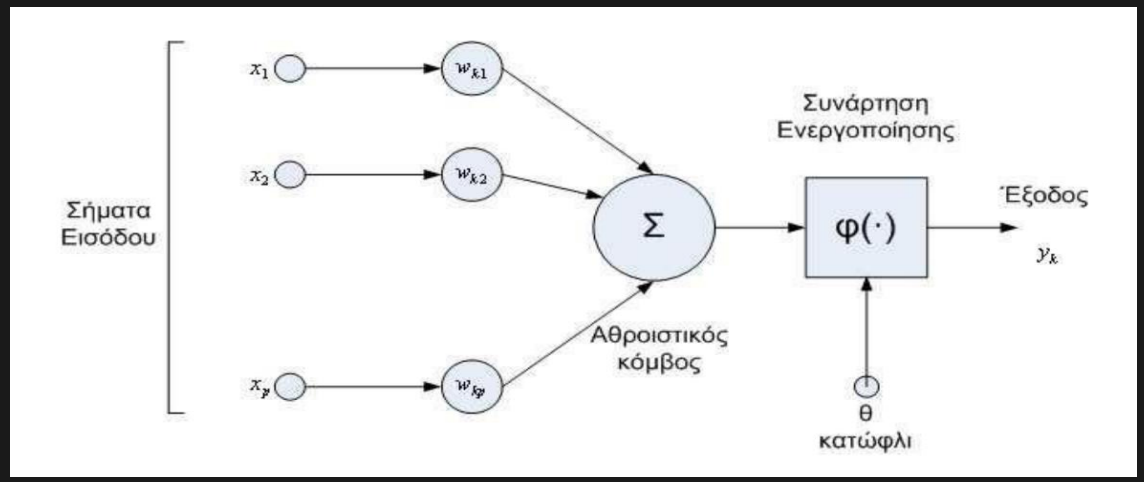
ΠΗΓΗ: http://repfiles.kallipos.gr/html_books/93/04a-main.html



ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

- ✓ **Σήματα εισόδου - συνεχείς μεταβλητές**, καθεμία από τις οποίες καλείται τιμή βάρους (weight).
- ✓ **Αθροιστής** (sum) προσθέτει τα σήματα εισόδου που έχουν μεταβληθεί από τις τιμές βάρους και παράγει την ποσότητα S
- ✓ **Συνάρτηση ενεργοποίησης** (activation function) διαμορφώνει την τελική τιμή του σήματος εξόδου y, σε συνάρτηση με την ποσότητα S και την τιμή κατωφλίου της συνάρτησης ενεργοποίησης.
- ✓ Μπορεί να έχει κάποιο **βάρος w0 - πόλωση** (bias) - ένα εξωτερικό ερέθισμα του νευρώνα το οποίο προστίθεται μαζί με τα υπόλοιπα σήματα εισόδου.

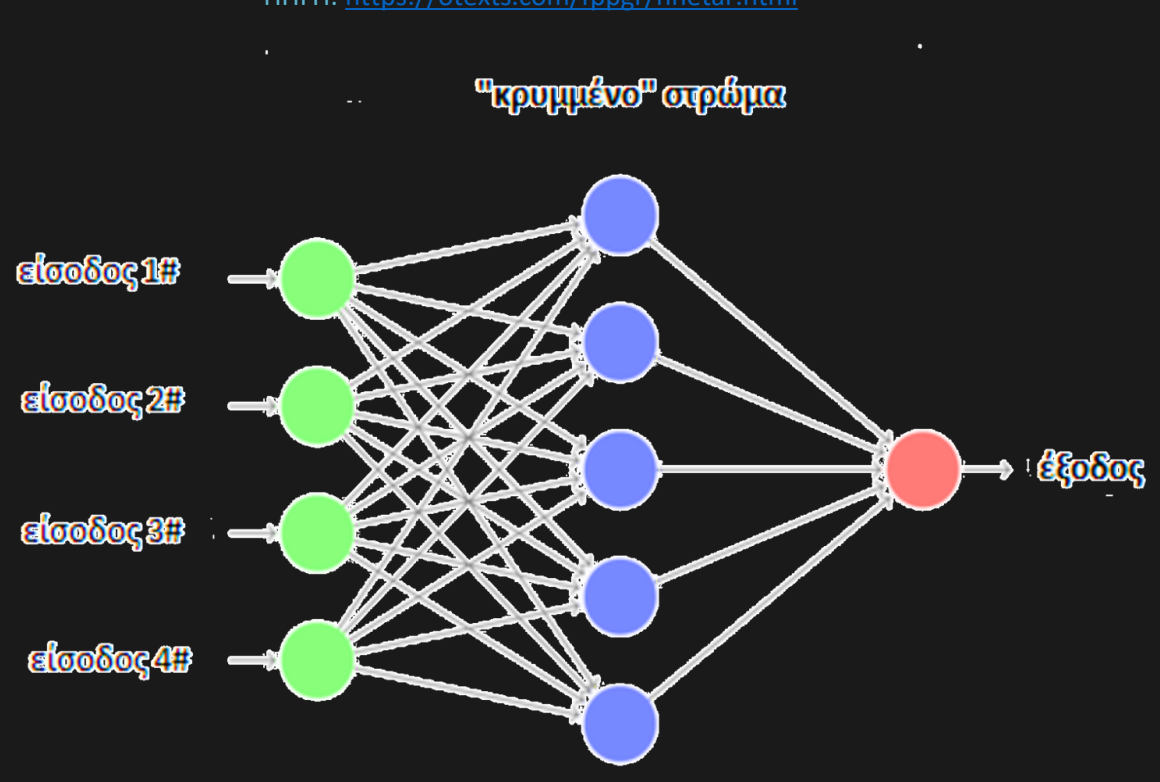
Η μορφή του τεχνητού νευρώνα (Amit,1989)



ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

- ✓ Τα τεχνητά νευρωνικά δίκτυα (artificial neural networks) αποτελούνται από τεχνητούς νευρώνες που είναι οργανωμένοι σε μία σειρά από στρώματα ή επίπεδα (layers). Τα επίπεδα αποτελούνται από μονάδες (units) ή κόμβους (nodes) που συνδέονται μεταξύ τους, ώστε να υπάρχουν διάφορες συνδέσεις μεταξύ των μονάδων των διαφόρων επιπέδων.
- ✓ Πλήρως συνδεδεμένοι (fully connected) όταν υπάρχει σύνδεση μεταξύ όλων των νευρώνων
- ✓ Σε κάθε άλλη περίπτωση οι νευρώνες είναι μερικώς συνδεδεμένοι (partially connected)

ΠΗΓΗ: https://otexts.com/fppgr/nnetar.html

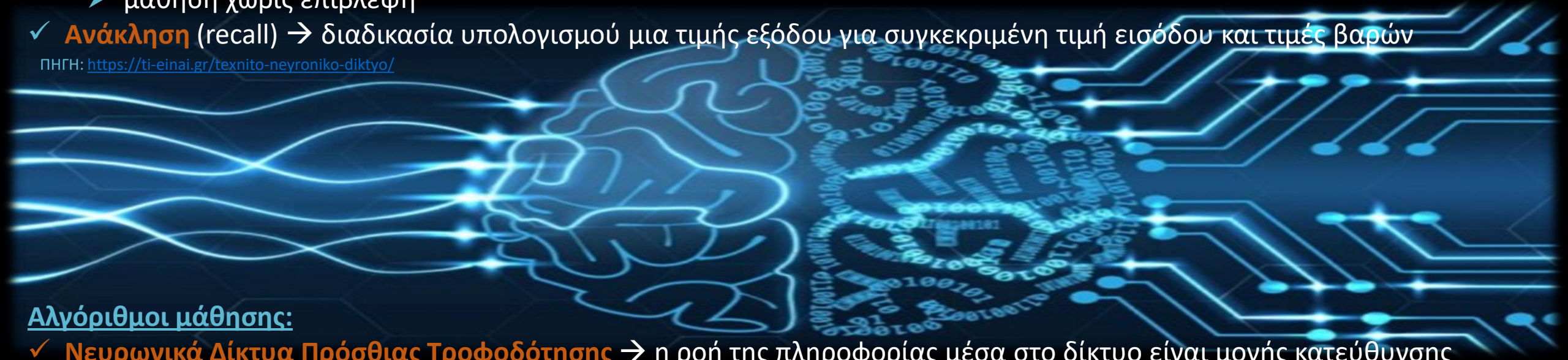


ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Βασικές μέθοδοι εκπαίδευσης:

- ✓ **Μάθηση** (learning) ή αλλιώς εκπαίδευση (training) → μια συγκεκριμένη τιμή εισόδου έχει ως αποτέλεσμα μια συγκεκριμένη τιμή εξόδου και έτσι μεταβάλλονται οι τιμές των βαρών
  - μάθηση με επίβλεψη
  - ενισχυτική μάθηση
  - μάθηση χωρίς επίβλεψη
- ✓ **Ανάκληση** (recall) → διαδικασία υπολογισμού μια τιμής εξόδου για συγκεκριμένη τιμή εισόδου και τιμές βαρών

ΠΗΓΗ: https://ti-einai.gr/texnito-neyroniko-diktyo/



Αλγόριθμοι μάθησης:

- ✓ **Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης** → η ροή της πληροφορίας μέσα στο δίκτυο είναι μονής κατεύθυνσης
- ✓ **Perceptron** → δίκτυο πρόσθιας τροφοδότησης χωρίς κρυφά επίπεδα
- ✓ **Κανόνας Δέλτα** → παραλλαγή του κανόνα του Perceptron που δεν εφαρμόζεται σε δίκτυα με κρυφά επίπεδα
- ✓ **Ανάστροφη Μετάδοση Λάθους** (back propagation) → οι νευρώνες στο επίπεδο εισόδου παράγουν κάποιο αποτέλεσμα, το οποίο αποτελεί είσοδο για το επόμενο επίπεδο (forward pass). Για να περιοριστεί το σφάλμα στην έξοδο οι τιμές βαρών αναπροσαρμόζονται και αυτό ονομάζεται «ανάστροφο πέρασμα» (backward pass) ή «ανάστροφη μετάδοση» (back propagation).

ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

ΣΥΝΕΛΙΚΤΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ (CONVOLUTIONAL NEURAL NETWORKS – CNN - CONVNET))

- ✓ Μάθηση με επίβλεψη
- ✓ Εντοπισμός και αναγνώριση προτύπων σε εικόνες
- ✓ Αποτελούνται από πολλά επίπεδα
- ✓ Εισάγονται ακατέργαστα διανύσματα εικόνων ως δεδομένα εισόδου και το επίπεδο εισόδου περιέχει τις τιμές των «pixel» κάθε εικόνας
- ✓ Εξάγονται χαρακτηριστικά κατά την ταξινόμηση εικόνων
- ✓ Τρόποι ταξινόμησης αντικειμένων:
  - ➢ εκπαίδευση από την αρχή
  - ➢ μεταφορά μάθησης
  - ➢ εξαγωγή χαρακτηριστικών

ΑΥΤΟΜΑΤΟΣ ΚΩΔΙΚΟΠΟΙΗΤΗΣ (AUTOENCODER)

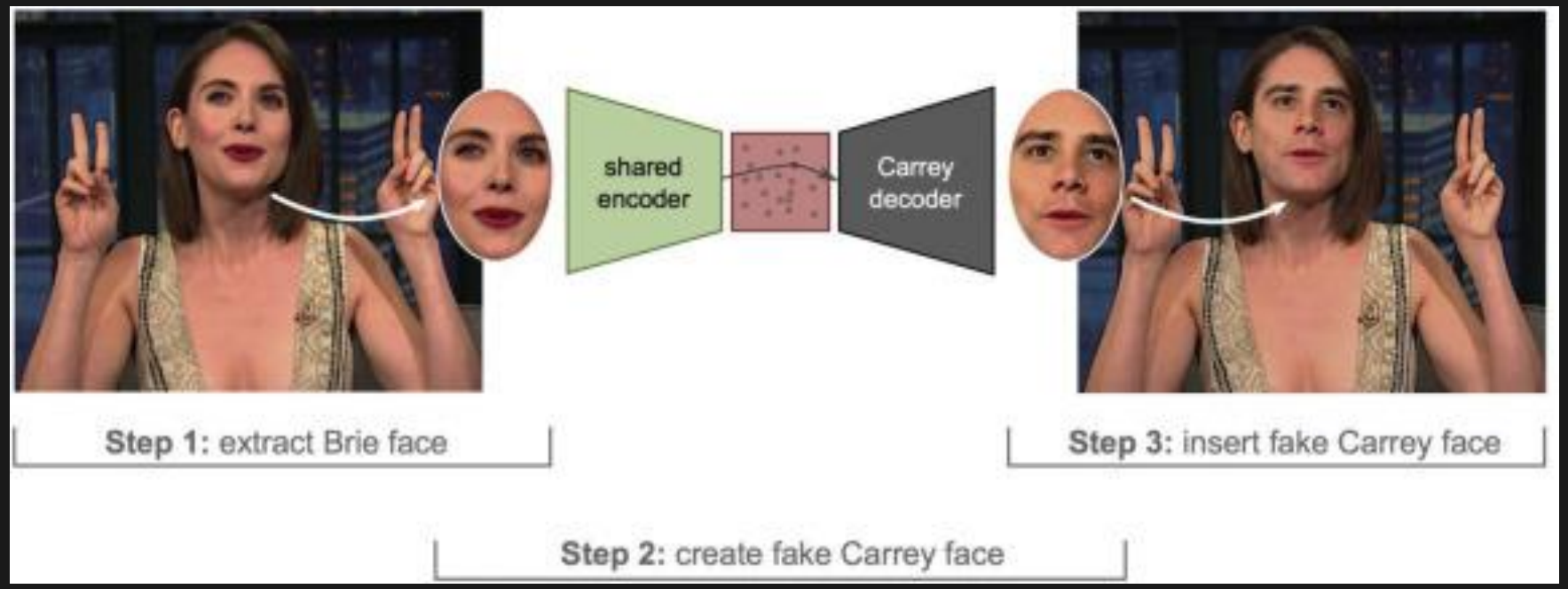
Νευρώνες που αποτελούνται από:

- ✓ Κωδικοποιητή, μειώνει τα μεγέθη εισόδου και συμπιέζει τα δεδομένα εισόδου σε μια κωδικοποιημένη αναπαράσταση
- ✓ Το εμπόδιο (Bottleneck) – οι χαμηλότερες δυνατές διαστάσεις των δεδομένων εισόδου
- ✓ Αποκωδικοποιητή, ανακατασκευάζει τα δεδομένα από την κωδικοποιημένη αναπαράσταση ώστε να είναι όσο το δυνατόν πιο κοντά στην αρχική είσοδο
- ✓ Απώλεια ανακατασκευής, η οποία είναι η μέθοδος που μετράει πόσο καλά αποδίδει ο αποκωδικοποιητής και πόσο κοντά είναι η έξοδος στην αρχική είσοδο

ΓΕΝΕΤΙΚΑ ΑΝΤΙΠΑΡΑΘΕΤΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ (GENERATIVE ADVERSARIAL NEURAL NETWORK - GAN)

- ✓ Δύο νευρωνικά δίκτυα (το δίκτυο δημιουργίας και το δίκτυο διακρίσεων) που ανταγωνίζονται μεταξύ τους όπως σε ένα παιχνίδι.
- ✓ Ο γεννήτορας τροφοδοτείται με τυχαίες εικόνες στην είσοδο και παράγει μία νέα εικόνα.
- ✓ Η νέα εικόνα καθώς και οι αληθινές εικόνες που εισήχθησαν τροφοδοτούν τον διευκρινιστή που προβλέπει αν η παραγόμενη εικόνα είναι αληθινή ή ψεύτικη.
- ✓ Διπλή ανατροφοδότηση, ο γεννήτορας ανατροφοδοτείται με τον διευκρινιστή και αυτός με αληθινές εικόνες για να βελτιώνεται και να μην μπορεί να ξεγελαστεί από τον γεννήτορα.

ΒΑΣΙΚΟΣ ΤΡΟΠΟΣ ΥΛΟΠΟΙΗΣΗΣ DEEPFAKE



ΠΗΓΗ: J. Kietzmann, L.W. Lee, I.P. McCarthy, T.C. Kietzmann, Deepfakes: Trick or treat?, Business Horizons, 63 (2) (2020), pp. 135-146, https://www.sciencedirect.com/science/article/abs/pii/S0007681319301600

Βήμα 1: Η περιοχή της εικόνας που δείχνει το πρόσωπο της Alison Brie εξάγεται από ένα αρχικό καρέ του βίντεο. Αυτή η εικόνα χρησιμοποιείται στη συνέχεια ως είσοδος σε ένα βαθύ νευρωνικό δίκτυο (DNN), μια τεχνική από τον τομέα της μηχανικής μάθησης και της τεχνητής νοημοσύνης.

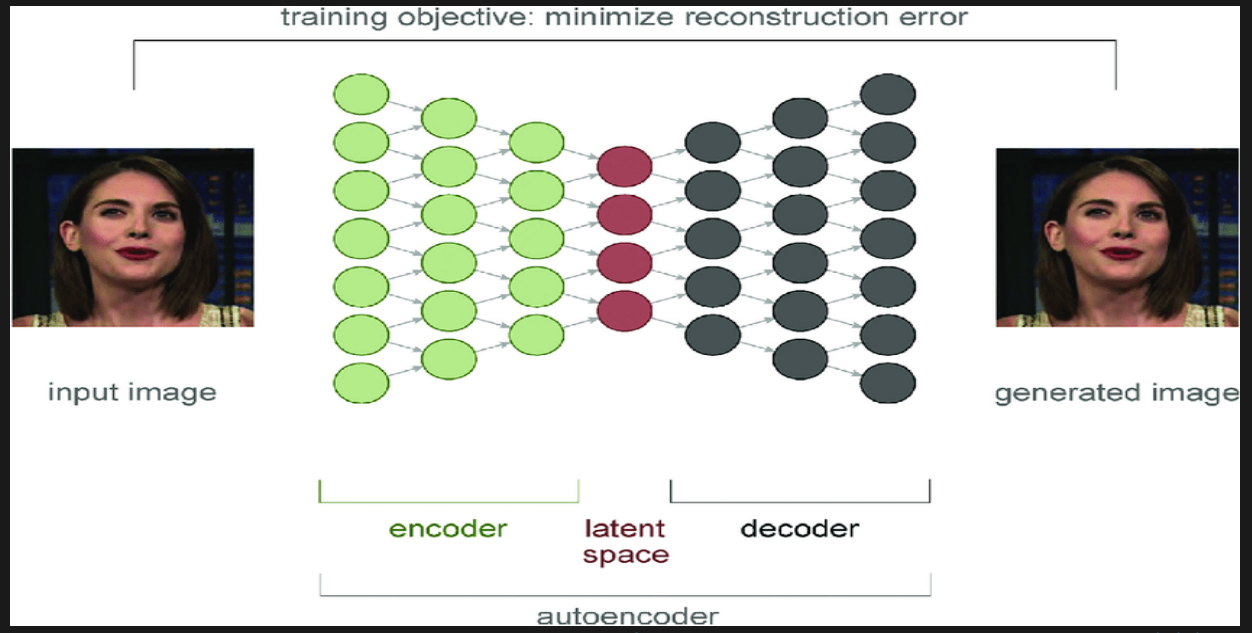
Βήμα 2: Το DNN παράγει αυτόματα μια εικόνα που ταιριάζει με την εικόνα του Jim Carrey αντί της Brie.

Βήμα 3: Αυτό το πρόσωπο που δημιουργείται εισάγεται στην αρχική εικόνα αναφοράς για να δημιουργηθεί το deepfake.

ΒΑΣΙΚΟΣ ΤΡΟΠΟΣ ΥΛΟΠΟΙΗΣΗΣ DEEPFAKE

- ✓ Τα deepfakes δημιουργούνται συνήθως με τη χρήση της αρχιτεκτονικής βαθιού νευρωνικού δικτύου, αυτή του **αυτόματου κωδικοποιητή (autoencoder)**.
- ✓ Οι αυτοκωδικοποιητές εκπαιδεύονται να αναγνωρίζουν βασικά χαρακτηριστικά μιας εικόνας εισόδου για να την αναδημιουργήσουν στη συνέχεια ως έξοδο.
- ✓ Κατά τη διαδικασία αυτή, το δίκτυο εκτελεί μεγάλη συμπίεση δεδομένων.
- ✓ Οι αυτόματοι κωδικοποιητές αποτελούνται από τρία επιμέρους τμήματα:

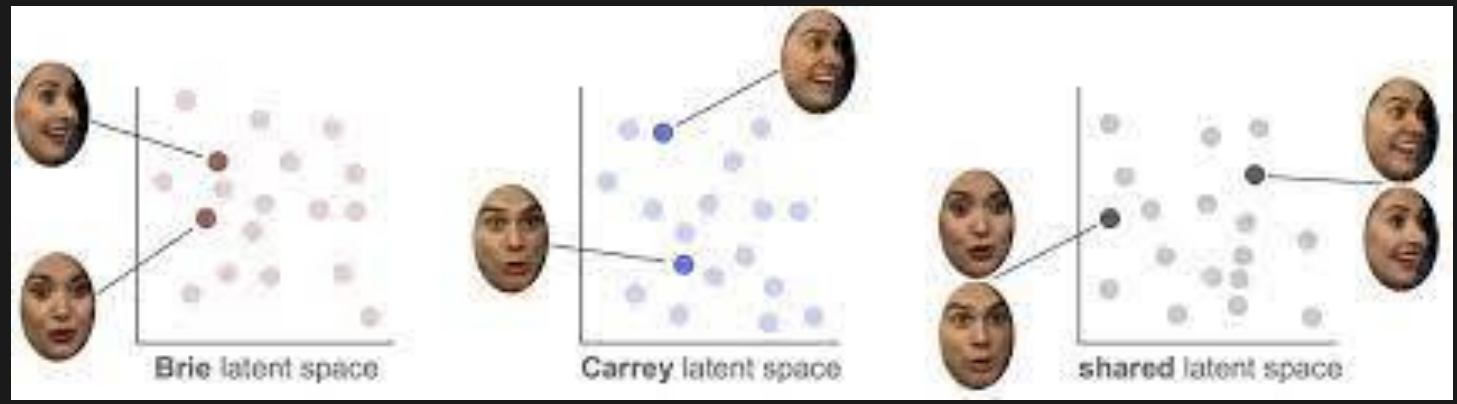
  - έναν **κωδικοποιητή - encoder** (που αναγνωρίζει τα βασικά χαρακτηριστικά ενός προσώπου εισόδου και συμπιέζει τα εικονοστοιχεία – pixel της εικόνας)
  - έναν **λανθάνοντα χώρο – latent space** (που αναπαριστά το πρόσωπο ως συμπιεσμένη έκδοση)
  - έναν **αποκωδικοποιητή – decoder** (αποσυμπίεση των πληροφοριών και ανακατασκευή της εικόνας εισόδου με όλες τις λεπτομέρειες όσο το δυνατόν πιο τέλεια)



ΠΗΓΗ: J. Kietzmann, L.W. Lee, I.P. McCarthy, T.C. Kietzmann, Deepfakes: Trick or treat?, Business Horizons, 63 (2) (2020), pp. 135-146, https://www.sciencedirect.com/science/article/abs/pii/S0007681319301600

ΒΑΣΙΚΟΣ ΤΡΟΠΟΣ ΥΛΟΠΟΙΗΣΗΣ DEEPFAKE

- ✓ Για την ανταλλαγή προσώπων χρειάζονται **δύο ζεύγη αυτόματων κωδικοποιητών**.
- ✓ Κάθε ζεύγος χρησιμοποιείται για εκπαίδευση σε ένα σύνολο εικόνων.
- ✓ Τα δύο ζεύγη έχουν **τον ίδιο κωδικοποιητή**.
- ✓ Η χρήση του ίδιου κωδικοποιητή και, ως εκ τούτου, η αναπαράσταση του λανθάνοντος χώρου για εικόνες δύο διαφορετικών ανθρώπων είναι **το κλειδί για την κατανόηση των deepfakes**.
- ✓ Αυτή η στρατηγική επιτρέπει στον κοινό κωδικοποιητή να βρίσκει και να μαθαίνει την **ομοιότητα μεταξύ δύο συνόλων εικόνων προσώπου**, οι οποίες είναι σχετικά απλές, επειδή τα πρόσωπα έχουν συνήθως παρόμοια χαρακτηριστικά, όπως τα μάτια, η μύτη, η θέση του στόματος.
- ✓ Εάν οι δύο αυτόματοι κωδικοποιητές εκπαιδεύονταν χωριστά, οι λανθάνουσες περιοχές δεν θα ευθυγραμμίζονταν (λανθάνουσα περιοχή Brie και Carrey παρακάτω). Η κοινή χρήση κωδικοποιητή θα οδηγήσει σε έναν **ευθυγραμμισμένο λανθάνοντα χώρο** (γκρίζες κουκκίδες). Οι αυτόματοι κωδικοποιητές μπορούν στη συνέχεια να χρησιμοποιηθούν για την **αντιστοίχιση ενός ατόμου με ένα άλλο άτομο**.



ΠΗΓΗ: J. Kietzmann, L.W. Lee, I.P. McCarthy, T.C. Kietzmann, Deepfakes: Trick or treat?, Business Horizons, 63 (2) (2020), pp. 135-146, https://www.sciencedirect.com/science/article/abs/pii/S0007681319301600

ΤΡΟΠΟΙ ΚΑΤΑΣΚΕΥΗΣ ΕΝΌΣ DEEPFAKE

Deepfake Φωτογραφία

- Ανταλλαγή προσώπου και σώματος → Πραγματοποίηση αλλαγών σε ένα πρόσωπο αντικαθιστώντας ή αναμειγνύοντας το πρόσωπο (ή το σώμα) με το πρόσωπο ή το σώμα κάποιου άλλου

Deepfake Ήχος

- Ανταλλαγή φωνής → Αλλαγή φωνής ή μίμηση της φωνής κάποιου άλλου
- Μετατροπή κειμένου σε ομιλία → Αλλαγή του ήχου σε μια ηχογράφηση πληκτρολογώντας νέο κείμενο

Deepfake Βίντεο

- Ανταλλαγή προσώπων → Αντικατάσταση του προσώπου κάποιου σε ένα βίντεο με το πρόσωπο κάποιου άλλου.
- Μορφοποίηση Προσώπου → Ένα πρόσωπο μεταμορφώνεται σε άλλο πρόσωπο μέσω μιας απρόσκοπτης μετάβασης
- «Κουκλοθέατρο» πλήρους σώματος → Μεταφορά της κίνησης από το σώμα ενός ατόμου στο σώμα ενός άλλου.

Εφαρμογές – Πλατφόρμες Δημιουργίας

DeepFaceLab, Face Swap, Zao, FakeApp, Cupace, Social media όπως Snapchat, Tik Tok και Instagram, DeepNude

Τομείς με θετικές επιδράσεις

Διαφήμιση, Μόδα, Κινηματογράφος, Βιομηχανία ψυχαγωγίας, Ιατρική

ΠΗΓΗ: https://petapixel.com/2022/07/22/megaportraits-high-res-deepfakes-created-from-a-single-photo/

Η ΣΚΟΤΕΙΝΗ ΠΛΕΥΡΑ ΤΩΝ DEEPFAKE

- ✓ Τα deepfakes έχουν γίνει «viral» και καθημερινό φαινόμενο παγκοσμίως και η χρήση τους γεννά νομικά, ηθικά και κοινωνικά διλήμματα και ερωτήματα.
- ✓ Λόγω του ανοιχτού κώδικα του αλγορίθμου, τα deepfakes μπορούν να αποτελέσουν όπλο οποιουδήποτε κακόβουλου χρήστη του διαδικτύου, αφού μπορεί να τα χρησιμοποιήσει ο κάθε χρήστης με πρόσβαση σε υπολογιστή συνδεδεμένο στο διαδίκτυο, και χωρίς να έχει στοιχειώδεις γνώσεις τεχνητής νοημοσύνης.
- ✓ Τα deepfakes έχουν ψηφιακή διάσταση.
- ✓ Μπορούν να αποτελέσουν απειλή για την κοινωνική ευταξία, για την ασφάλεια της πληροφορίας, αλλά και για έννομα αγαθά και ελευθερίες των πολιτών σε παγκόσμιο επίπεδο.

ΠΗΓΗ: FORBES https://www.forbes.com/sites/petersuciu/2020/12/11/deepfake-star-wars-videos-portent-ways-the-technology-could-be-employed-for-good-and-bad/

ΤΟ DEEPFAKE ΩΣ ΚΥΒΕΡΝΟΕΓΚΛΗΜΑ

- ✓ Υπ' αρ. 185 Σύμβαση του Συμβουλίου της Ευρώπης (Βουδαπέστη, 23.11.2001) για την προστασία της κοινωνίας από το κυβερνοέγκλημα.

- ✓ Ν 4411/2016 (ΦΕΚ Α΄ 142/3.8.2016)

- ✓ e-crime, computer-crime, cybercrime και internet related crime, ηλεκτρονικό έγκλημα και διαδικτυακό έγκλημα και κυβερνοέγκλημα ή έγκλημα του κυβερνοχώρου

- ✓ Το ηλεκτρονικό έγκλημα μπορεί να χαρακτηριστεί

α) ως έγκλημα που διαπράττεται με τη χρήση ηλεκτρονικών υπολογιστών (computer crimes)

β) ως ένα ήδη υπάρχον έγκλημα με τη διαφορά ότι διαπράττεται με υπολογιστή και

γ) ως μια εγκληματική συμπεριφορά που εκδηλώνεται με την με οποιοδήποτε τρόπο συμμετοχή ενός ηλεκτρονικού υπολογιστή ή μέσω διαδικτύου (cyber crimes).

- ✓ Μπορεί να διακριθεί σε:

α) γνήσιο πληροφορικό έγκλημα, όταν διαπράττεται μέσω ηλεκτρονικού υπολογιστή και μέσω συστημάτων πληροφοριών (λ.χ. απάτη, πλαστογραφία),

β) έγκλημα με ψηφιακό περιεχόμενο, όταν διακινείται παράνομο περιεχόμενο μέσω συστημάτων πληροφοριών (λ.χ. παιδική πορνογραφία),

γ) έγκλημα κατά πληροφοριακών συστημάτων, όταν προσβάλλεται η εμπιστευτικότητα, ακεραιότητα και διαθεσιμότητα των πληροφοριακών συστημάτων και των ψηφιακών δεδομένων.

- ✓ Βασικό στοιχείο ο ηλεκτρονικός υπολογιστής με σύνδεση σε σύστημα πληροφοριών (ή στόχος της επίθεσης, ή το βασικό μέσο / εργαλείο της επίθεσης, ή το βοηθητικό μέσο / εργαλείο για την επίθεση.

- ✓ Έγκλημα που τελείται από απόσταση, αφού αλλού ενεργεί ο δράστης και αλλού επέρχεται το αξιόποινο αποτέλεσμα.

ΠΗΓΗ: https://www.semifind.gr/news/view/Nea/6172/Cyber-Crime---to-kyvernoegklima-apo-tin-proanakrisi-sto-akroatirio-%7C-seminario-me-to-manoli-sfakianaki

ΤΟ DEEPFAKE ΩΣ ΜΕΣΟ ΠΑΡΑΠΛΗΡΟΦΟΡΗΣΗΣ

- ✓ Μέσω των διάσημων πλατφορμών Tik Tok, Twitter, Instagram και Facebook εύκολα μια είδηση μπορεί να γίνει «viral».
- ✓ Ο καθένας μας μπορεί να γίνει ένας μικρός δημοσιογράφος που θα διασπείρει μια αληθινή αλλά και μια ψεύτικη είδηση.
- ✓ Δύναμη της εικόνας και εξάρτηση στα social media - Μετάβαση από την εικόνα στο βίντεο
- ✓ Η παραπληροφόρηση γνωστή ως **«fake news» ή «hoax»** και η διασπορά ψευδών ειδήσεων έχουν λάβει νέες διαστάσεις στην ψηφιακή εποχή και απασχολούν τα νομικά συστήματα σε εθνικό και διεθνές επίπεδο.
- ✓ Fake news είναι οι **ψευδείς ειδήσεις**, οι οποίες διαδίδονται από ειδησεογραφικές ιστοσελίδες, ψεύτικους λογαριασμούς προσώπων σε μέσα κοινωνικής δικτύωσης, καθώς και από τα λεγόμενα «troll».
- ✓ Πλήγμα για την δημοκρατία και την κοινωνία.

ΠΗΓΗ: https://www.skai.gr/news/technology/pos-tha-anagnoriso-ta-fake-news-praktikes-symvoules

ΠΗΓΗ: https://dribbble.com/shots/4747715-Fake-News-Exhibition-GIF

ΑΡΘΡΟ 191 ΠΚ

«Όποιος δημόσια ή μέσω του διαδικτύου διαδίδει ή διασπείρει με οποιονδήποτε τρόπο ψευδείς ειδήσεις με αποτέλεσμα να προκαλέσει φόβο σε αόριστο αριθμό ανθρώπων ή σε ορισμένο κύκλο ή κατηγορία προσώπων που αναγκάζονται έτσι να προβούν σε μη προγραμματισμένες πράξεις ή σε ματαίωσή τους, με κίνδυνο να προκληθεί ζημία στην οικονομία, στην αμυντική ικανότητα της χώρας ή στη δημόσια υγεία τιμωρείται με φυλάκιση έως τρία (3) έτη ή με χρηματική ποινή. Με την ίδια ποινή τιμωρείται και ο πραγματικός ιδιοκτήτης ή εκδότης του μέσου με το οποίο τελέστηκαν οι πράξεις του παρόντος.»

Αντικειμενική Υπόσταση:

- Εγκληματική συμπεριφορά → δημόσια ή μέσω του διαδικτύου διάδοση ή διασπορά με οποιονδήποτε τρόπο ψευδών ειδήσεων, με αποτέλεσμα να προκληθεί φόβος σε αόριστο αριθμό ανθρώπων ή σε ορισμένο κύκλο ή κατηγορία προσώπων.
- Είδηση → ανακοίνωση γεγονότος του παρόντος ή του πρόσφατου παρελθόντος και όχι του μέλλοντος
- Έγκλημα συγκεκριμένης διακινδύνευσης → πρόκληση φόβου ως αποτέλεσμα
- Αιτιώδη συνάφεια μεταξύ φόβου και πράξεων
- Προστατευόμενο έννομο αγαθό → η δημόσια τάξη τόσο στην κοινωνική όσο και στην πολιτειακή της μορφή
- Δράστης μπορεί να είναι οποιοσδήποτε
- Αντικείμενο του εγκλήματος → αόριστος αριθμός προσώπων που ειρηνικά συνυπάρχουν σε ορισμένο κοινωνικό χώρο σε κατάσταση ευταξίας αλλά και ορισμένος κύκλος προσώπων.

Υποκειμενική Υπόσταση:

Δόλος, τουλάχιστον ενδεχόμενος

Τιμώρηση:

Ο εκ δόλου δράστης τιμωρείται με φυλάκιση έως 3 έτη ή χρηματική ποινή.

ΠΑΡΑΔΕΙΓΜΑΤΑ ΠΑΡΑΠΛΗΡΟΦΟΡΗΣΗΣ ΜΕ ΤΗ ΧΡΗΣΗ DEEPFAKES

ΤΑ DEEPFAKES ΩΣ ΜΕΣΟ ΑΛΛΟΙΩΣΗΣ ΕΝΟΣ ΠΡΟΣΤΑΤΕΥΟΜΕΝΟΥ ΕΡΓΟΥ ΠΝΕΥΜΑΤΙΚΗΣ ΙΔΙΟΚΤΗΣΙΑΣ

- ✓ **Ως έργο νοείται το ανθρώπινο δημιούργημα**, χωρίς πάντως να αποκλείεται η χρήση τεχνικών μέσων.
- ✓ Το έργο πρέπει να είναι **πρωτότυπο** (αρ. 2 § 1 Ν. 2121/1993)
- ✓ **Δημιουργός** είναι ο αρχικός δικαιούχος του περιουσιακού και του ηθικού δικαιώματος επ' αυτού (αρ. 6 Ν. 2121/1993)
- ✓ **Περιουσιακό δικαίωμα** είναι η οικονομική εκμετάλλευση του έργου από τον δημιουργό, όπως να επιτρέπει ή να απαγορεύει: α) Την εγγραφή και την άμεση ή έμμεση, προσωρινή ή μόνιμη αναπαραγωγή των έργων τους με οποιοδήποτε μέσο και μορφή, εν όλω ή εν μέρει β) Τη μετάφραση των έργων τους γ) Τη διασκευή, την προσαρμογή ή άλλες μετατροπές των έργων τους κ.α. (αρ. 3 ν. 2121/1993)
- ✓ **Ηθικό δικαίωμα** είναι η προστασία του προσωπικού δεσμού του δημιουργού προς το έργο του (αρ. 1 ν. 2121/1993) όπως α) η απόφαση για το χρόνο, τον τόπο και τον τρόπο κατά τους οποίους το έργο θα γίνει προσιτό στο κοινό (δημοσίευση), β) η αναγνώριση πατρότητας πάνω στο έργο κ.α. (αρ. 4 ν. 2121/1993)
- ✓ **Περιορισμοί** του δικαιώματος της πνευματικής ιδιοκτησίας που διακρίνονται χάριν του γενικού συμφέροντος όπως η εξαίρεση υπέρ της ελευθερίας της έκφρασης κ.α. και χάριν των συμφερόντων ιδιωτών και συγκεκριμένα επί αναπαραγωγής για ιδιωτική χρήση (έλλειψη κοινού και απεύθυνσης σ' αυτό) (αρ. 18 – 28Γ ν. 2121/1993)
- ✓ **Συγγενικά δικαιώματα** είναι "τα μέσα μεταφοράς" του πνευματικού έργου στο κοινό, μορφή προσφοράς του έργου στο κοινό, πχ. απαγγελία, ερμηνεία κλπ (αρ. 46 - 51Α του ν. 2121/1993)
- ✓ **Προσβολή των πνευματικών δικαιωμάτων** → όταν τρίτος θίγει την άσκηση της εξουσίας του δικαιούχου, ήτοι προβαίνει σε συμπεριφορά στην οποία μόνον ο δικαιούχος νομιμοποιείται.
- ✓ Αστικές κυρώσεις (Άρθρο 65 ν. 2121/1993) - Διοικητικές κυρώσεις (Άρθρο 65Α ν. 2121/1993) - Ποινικές κυρώσεις (66 του ν. 2121/1993)

Δημοσίευση του αλλοιωμένου βίντεο:

- ✓ Στόχος του δημιουργού ενός αλλοιωμένου βίντεο είναι να το διανείμει στο κοινό μέσω των κοινωνικών δικτύων.
- ✓ Βασική προϋπόθεση για την παραβίαση δικαιώματος πνευματικής ιδιοκτησίας είναι η απουσία της συγκατάθεσης του δικαιούχου.
- ✓ Ορισμένα δικαιώματα που παραβιάζονται είναι το περιουσιακό δικαίωμα αναπαραγωγής (το οποίο εφαρμόζεται ακόμη και όταν αλλοιώνεται η μορφή του προστατευόμενου έργου), το περιουσιακό δικαίωμα παρουσίασης στο κοινό, το δικαίωμα διασκευής και το ηθικό δικαίωμα του δημιουργού.
- ✓ Ωστόσο, υπάρχουν εξαιρέσεις και δεν υπάρχει πάντα προσβολή των δικαιωμάτων πνευματικής ιδιοκτησίας με τη χρήση της τεχνολογίας deepfake, οι οποίες προβλέπονται κυρίως στο άρθρο 5 της Οδηγίας 2001/29, το οποίο επιτρέπει την χρήση ενός προστατευόμενου έργου «για γελοιογραφία, παρωδία ή μίμηση».

ΖΗΤΗΜΑΤΑ ΠΑΡΑΒΙΑΣΗΣ ΠΝΕΥΜΑΤΙΚΩΝ ΔΙΚΑΙΩΜΑΤΩΝ ΜΕ ΤΗ ΧΡΗΣΗ ΤΩΝ DEEPFAKES

Η χρήση της βιβλιοθήκης φωτογραφιών

- Είναι συμβατή η χρήση φωτογραφιών που εισάγονται ως δεδομένα εικόνας στον αλγόριθμο deepfake με την πνευματική ιδιοκτησία;
- Οδηγία 2019/790 → εξαίρεση σχετικά με την εξόρυξη δεδομένων (data mining) → δικαίωμα στον χρήστη τροφοδότησης της τεχνητής νοημοσύνης με φωτογραφίες προσώπων

Το δικαίωμα επί της ιδίας εικόνας

- Το Deepfake αποτελεί ξεκάθαρη απόδειξη ότι η εικόνα του προσώπου είναι αντικείμενο εκμετάλλευσης και εμπορευματοποίησης.
- Στην Αμερική → «δικαίωμα στην δημοσιότητα» (right to publicity)
- Ως εικόνα νοείται η εξωτερική μορφή του προσώπου, η οποία το εξατομικεύει, ενώ το right of publicity έγκειται στην εξουσία οικονομικής εκμετάλλευσης στοιχείων της προσωπικότητας, όπως η εικόνα.
- Στην Γερμανία → μορφή διανοητικής ιδιοκτησίας ήδη από το 1907.
- Στην Ελλάδα → προστασία της προσωπικότητας.

- Νομολογία ΕΔΔΑ → Απόφαση Von Hannover κατά Γερμανίας το 2004 → δικαίωμα επί της ίδιας εικόνας ακόμα και ενός δημοσίου προσώπου → Απόφαση του έτους 2012 → αναλογικότητα μεταξύ ελευθερίας έκφρασης και σεβασμού του ιδιωτικού βίου.
- Η συγκατάθεση του προσώπου ή των αντιπροσώπων του υπαναχωρεί μπροστά στο δημόσιο συμφέρον του κοινού για ενημέρωση σχετικά με δημόσια πρόσωπα.
- Η ελευθερία έκφρασης υπερτερεί κατά τεκμήριο του δικαιώματος επί της ίδιας εικόνας. Όμως, το αντίθετο για μη δημόσια πρόσωπα

ΠΗΓΗ: https://lawadvocategroup.com/right-of-publicity/

ΤΑ DEEPFAKES ΩΣ ΜΕΣΟ ΠΡΟΣΒΟΛΗΣ ΠΡΟΣΩΠΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

- ✓ Ως **δεδομένα προσωπικού χαρακτήρα** νοείται «κάθε πληροφορία που αφορά ταυτοποιημένο ή ταυτοποιήσιμο φυσικό πρόσωπο («υποκείμενο των δεδομένων») (άρθρο 4 περ. 1 ΓΚΠΔ)
- ✓ **Απλά και ευαίσθητα** προσωπικά δεδομένα
- ✓ Ως **επεξεργασία** νοείται «κάθε πράξη ή σειρά πράξεων που πραγματοποιείται με ή χωρίς τη χρήση αυτοματοποιημένων μέσων, σε δεδομένα προσωπικού χαρακτήρα ή σε σύνολα δεδομένων προσωπικού χαρακτήρα, όπως η συλλογή, η καταχώριση, η οργάνωση, η διάρθρωση, η αποθήκευση, η προσαρμογή ή η μεταβολή, η ανάκτηση, η αναζήτηση πληροφοριών, η χρήση, η κοινολόγηση με διαβίβαση, η διάδοση ή κάθε άλλη μορφή διάθεσης, η συσχέτιση ή ο συνδυασμός, ο περιορισμός, η διαγραφή ή η καταστροφή». (αρ. 4 περ. 2 ΓΚΠΔ)
- ✓ **Νόμιμος λόγος επεξεργασίας** (αρ. 5 παρ. 1 περ. α' ΓΚΠΔ)
- ✓ **Συναίνεση** του υποκειμένου σε επεξεργασία (αρ. 6 ΓΚΠΔ) ελεύθερη και ελευθέρως ανακληθείσα (αρ. 7 ΓΚΠΔ)
- ✓ Αρχή της **αναλογικότητας** → σύνδεση σκοπού επεξεργασίας με δεδομένα, όσο το δυνατόν λιγότερα
- ✓ Αρχή της **καθορισμένης χρονικής διάρκειας** διατήρησης των δεδομένων
- ✓ Το **δικαίωμα στη λήθη** → το υποκείμενο των δεδομένων ζητά τη διαγραφή των προσωπικών του δεδομένων, εφόσον δεν επιθυμεί πια αυτά τα δεδομένα να αποτελούν αντικείμενο επεξεργασίας.
- ✓ **Αξιόποινες συμπεριφορές** είναι παραβάσεις:

α) των υποχρεώσεων προστασίας των προσωπικών δεδομένων,

β) λόγω μη συμμόρφωσης με αποφάσεις της Αρχής και

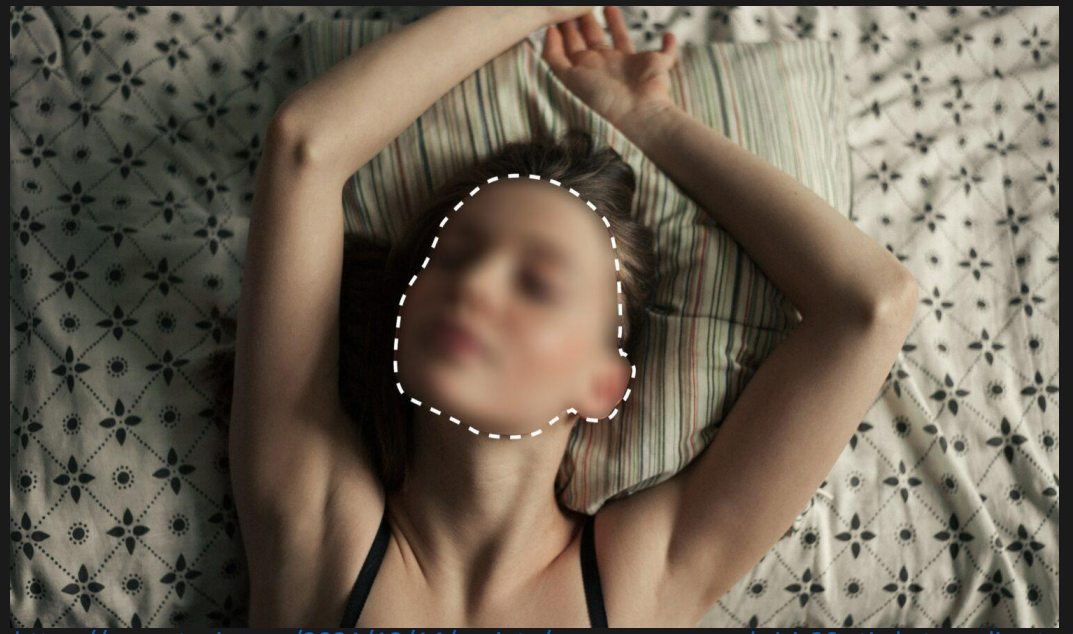
γ) που αφορούν τη χωρίς δικαίωμα επέμβαση σε αρχείο προσωπικών δεδομένων και την τέλεση περαιτέρω πράξεων που προσβάλλουν τα προσωπικά δεδομένα.

- ✓ **Άρθρο 38 του νέου Νόμου 4624/2019** → «Όποιος, χωρίς δικαίωμα: α) επεμβαίνει με οποιονδήποτε τρόπο σε σύστημα αρχειοθέτησης δεδομένων προσωπικού χαρακτήρα, και με την πράξη του αυτή λαμβάνει γνώση των δεδομένων αυτών· β) τα αντιγράφει, αφαιρεί, αλλοιώνει, βλάπτει, συλλέγει, καταχωρεί, οργανώνει, διαρθρώνει, αποθηκεύει, προσαρμόζει, μεταβάλλει, ανακτά, αναζητεί πληροφορίες, συσχετίζει, συνδυάζει, περιορίζει, διαγράφει, καταστρέφει, τιμωρείται με φυλάκιση μέχρι ενός (1) έτους, εάν η πράξη δεν τιμωρείται βαρύτερα με άλλη διάταξη»

ΠΗΓΗ: https://www.epam.com/insights/blogs/personal-data-protection-what-does-it-really-mean-today

DEEP FAKE PORNOGRAPHY

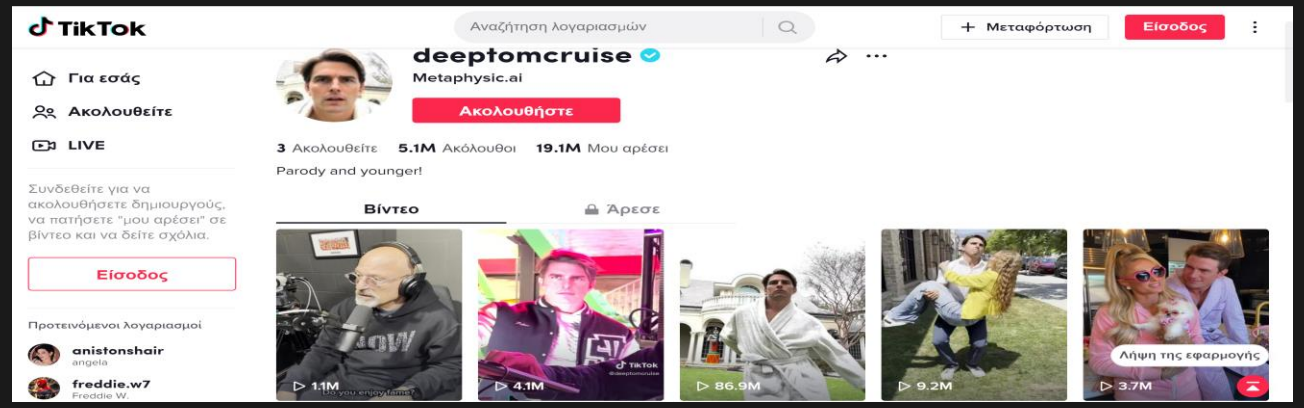
- ✓ **«Εκδικητική πορνογραφία», «Μη συναινετική πορνογραφία», «Σεξουαλική κακοποίηση μέσω εικόνας» = δημοσιοποίηση στο διαδίκτυο φωτογραφιών ή βίντεο με σεξουαλικό περιεχόμενο χωρίς τη συναίνεση του εικονιζόμενου προσώπου = Σεξουαλικό έγκλημα = «διαδικτυακός βιασμός» («cyber rape») = σεξουαλική βία.**
- ✓ **Μορφές:** «sexting», «upskirting», «downblousing», μέσω hacking, «sextortion», «pornographic photoshopping» και «deepfake videos»
- ✓ Παραβίαση της σεξουαλικής ιδιωτικής ζωής, σεξουαλική ταπείνωση και εκμετάλλευση της σωματικής, ψυχικής ή οικονομικής κακοποίησης ατόμων
- ✓ Γερμανία → «εκδικητική πορνογραφία» ως sui generis έγκλημα
- ✓ Αγγλία → ποινικοποίηση του revenge porn το 2015.
- ✓ Αγγλία → ποινικοποίηση του deepfake porn 12/2022
- ✓ Κίνα → 1η Ιανουαρίου 2020 κυβερνητική πολιτική κατά deepfakes
- ✓ Ελλάδα → εφαρμογή ποινικών διατάξεων του νόμου «για την προστασία των δεδομένων προσωπικού χαρακτήρα» (Ν. 4624/2019)



ΠΗΓΗ: https://www.tovima.gr/2021/12/14/society/revenge-porn-ayksisi-66-stis-kataggelies-to-2021/

ΠΗΓΗ: https://www.vox.com/2018/1/31/16932264/reddit-celebrity-porn-face-swapping-dystopia

ΜΟΡΦΗ ΣΑΤΙΡΑΣ



DEAD OR ALIVE ?

ΤΑ DEEPFAKES ΩΣ ΜΕΣΟ ΔΥΣΦΗΜΗΣΗΣ

- ✓ Η τεχνολογία deepfake μπορεί να αποτελέσει έγκλημα κατά της τιμής, αφού βασίζεται σε ένα ψέμα, διότι με την ανταλλαγή προσώπου υπονοείται ότι το άτομο συμμετείχε με τη βούλησή του στο βίντεο, ενώ στην πραγματικότητα αυτό δεν συνέβη ποτέ.

- ✓ Αρκεί να αποδειχθεί ότι η εικόνα ή το βίντεο προορίζεται να αποτελέσει δήλωση γεγονότος

- ✓ **Άρθρο 362 ΠΚ** → **Δυσφήμηση** «Όποιος με οποιονδήποτε τρόπο ενώπιον τρίτου ισχυρίζεται ή διαδίδει για κάποιον άλλον γεγονός που μπορεί να βλάψει την τιμή ή την υπόληψή του τιμωρείται με φυλάκιση έως ένα έτος ή χρηματική ποινή. Αν η πράξη τελέστηκε δημόσια με οποιονδήποτε τρόπο ή μέσω διαδικτύου, επιβάλλεται φυλάκιση έως τρία έτη ή χρηματική ποινή.»

- ✓ **Άρθρο 363 ΠΚ** → **Συκοφαντική Δυσφήμηση** «Αν στην περίπτωση του προηγούμενου άρθρου, το γεγονός είναι ψευδές και ο υπαίτιος γνώριζε ότι αυτό είναι ψευδές τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών και χρηματική ποινή και αν τελεί την πράξη δημόσια με οποιονδήποτε τρόπο ή μέσω του διαδικτύου, με φυλάκιση τουλάχιστον έξι μηνών και χρηματική ποινή.»

- ✓ Δυσφήμηση → **ισχυρισμός** (ανακοίνωση του δράστη) ή **διάδοση** (μετάδοση ανακοίνωσης που έγινε από άλλον) γεγονότος ενώπιον τρίτου → που μπορεί να **βλάψει την τιμή** και την υπόληψή του → την εκτίμηση που απολαμβάνει το άτομο στην κοινωνία

- ✓ **Υποκειμενική υπόσταση:** δόλος

- ✓ **Άρθρο 367 ΠΚ** → εξαιρέσεις από τον άδικο χαρακτήρα της πράξης όταν υπάρχει **δικαιολογημένο ενδιαφέρον**

ΠΗΓΗ: https://blog.ipleaders.in/defamation-law-in-india/

ΤΟ DEEPFAKE ΩΣ ΡΑΤΣΙΣΤΙΚΟ ΚΑΙ ΞΕΝΟΦΟΒΙΚΟ ΥΛΙΚΟ

- ✓ Άρθρο 2 παρ. 1 του Πρόσθετου Πρωτοκόλλου της Σύμβασης της Βουδαπέστης «κάθε γραπτό υλικό, εικόνα ή άλλη έκφραση ιδεών ή θεωριών που υποστηρίζουν, προάγουν ή υποδαυλίζουν το μίσος τις διακρίσεις ή την βία κατά κάποιου ατόμου ή ομάδας ατόμων με βάση την φυλή, το χρώμα, την καταγωγή, την εθνική ή την εθνοτική προέλευση, καθώς και την θρησκεία, εάν αυτή χρησιμοποιείται ως πρόσχημα για κάποιον από τους ανωτέρω παράγοντες»

- ✓ Θύματα των εν λόγω εγκλημάτων μίσους είναι πρόσωπα που έχουν ως χαρακτηριστικό ότι ανήκουν σε μια εξατομικευμένη ομάδα που διακρίνεται με βάση κάποιο χαρακτηριστικό, όπως το χρώμα, την καταγωγή ή τη θρησκεία.

- ✓ Ν 927/1979 → Ν 4285/2014 → διακρίσεις, μίσος ή βία κατά προσώπου ή ομάδας προσώπων με ιδιαίτερα χαρακτηριστικά → α) κίνδυνος και απειλή για τη δημόσια τάξη, τη ζωή, την ελευθερία ή τη σωματική ακεραιότητα των ως άνω προσώπων και β) διάπραξη φθοράς ή βλάβης πραγμάτων που χρησιμοποιούνταν από τις παραπάνω ομάδες ή πρόσωπα.

- ✓ 81Α ΠΚ → Έγκλημα με ρατσιστικά χαρακτηριστικά

- ✓ Τα deepfakes ως επί το πλείστον σχετίζονται με την πορνογραφία → κατά κύριο λόγο μια συνέπεια της ασέβειας προς τις γυναίκες και της αντικειμενοποίησης του γυναικείου σώματος

- ✓ Έμφυλη βία → Γυναικοκτονίες και Αυτοκτονίες, όπως η Λίνα Κοεμτζή

ΠΗΓΗ: https://www.efsyn.gr/ellada/koinonia/325335_i-emfyli-bia-toy-diadiktyoy-kolympa-sta-kena-toy-nomoy

ΤΟ DEEPFAKE ΩΣ ΜΕΣΟ ΕΚΦΟΒΙΣΜΟΥ - ΕΚΒΙΑΣΜΟΥ

- ✓ **Άρθρο 385 ΠΚ** «1. Όποιος, εκτός από τις περιπτώσεις του άρθρου 380, με σκοπό να αποκομίσει ο ίδιος ή άλλος παράνομο περιουσιακό όφελος, εξαναγκάζει κάποιον με βία ή απειλή σε πράξη, παράλειψη ή ανοχή από την οποία επέρχεται ζημία στην περιουσία του εξαναγκαζομένου ή άλλου τιμωρείται με φυλάκιση τουλάχιστον ενός έτους και χρηματική ποινή…»

- ✓ **Αντικειμενική Υπόσταση:** Εξαναγκασμός κάποιου με βία ή με απειλή → επιβολή συμπεριφοράς μη ηθελημένης από τον παθόντα

Περιουσιακή ζημία → να προκληθεί διά της εκβίασης

- ✓ **Υποκειμενική υπόσταση:** δόλος και σκοπό προσπορισμού περιουσιακού οφέλους (υπερχειλής υποκειμενική υπόσταση)

- ✓ **Cyberbullying** → διαδικτυακός εκφοβισμός είναι οποιαδήποτε επαναλαμβανόμενη πράξη εκφοβισμού, επιθετικότητας, παρενόχλησης, τρομοκρατικής ή αυταρχικής συμπεριφοράς που πραγματοποιείται μέσω της χρήσης ψηφιακών συσκευών.

- ✓ Η **σεξουαλική εκβίαση [sextortion]** είναι συχνό φαινόμενο στις μέρες μας και συντρέχει όταν ο δράστης εκβιάζει το θύμα με προσωπικό ή οικονομικό κίνητρο και πάλι με την δημοσίευση προσωπικών/σεξουαλικών εικόνων ή βίντεο.

- ✓ Πολύ συχνά μάλιστα μπορεί να πάρει την μορφή της **απάτης σεξουαλικής εκβίασης [sexrtotion scam],** όπου ο δράστης αποστέλλει μαζικά μηνύματα στο θύμα ότι δήθεν θα αποκαλύψει ευαίσθητο υλικό που το απεικονίζει εάν δεν καταθέσει το Χ οικονομικό ποσό σε τραπεζικό λογαριασμό που του αποστέλλει.

ΠΗΓΗ: https://ied.eu/project-updates/cyberbullying-what-is-the-situation-in-greece/

ΤΑ DEEPFAKES ΩΣ ΜΕΣΟ ΑΠΑΤΗΣ

Άρθρο 386 ΠΚ «1. Όποιος με την εν γνώσει παράσταση ψευδών γεγονότων σαν αληθινών ή την αθέμιτη απόκρυψη ή παρασιώπηση αληθινών γεγονότων βλάπτει ξένη περιουσία πείθοντας κάποιον σε πράξη, παράλειψη ή ανοχή με σκοπό από τη βλάβη αυτής της περιουσίας να αποκομίσει ο ίδιος ή άλλος παράνομο περιουσιακό όφελος τιμωρείται με φυλάκιση...».

Αντικειμενική υπόσταση:

α) πράξη εξαπάτησης,
β) πλάνη θύματος,
γ) περιουσιακή διάθεση του θύματος,
δ) βλάβη περιουσίας του θύματος ή άλλου, που πρέπει να αντιστοιχεί στο περιουσιακό όφελος που σκόπευε ο δράστης
ε) ο αντικειμενικός αιτιώδης σύνδεσμος μεταξύ όλων των προαναφερόμενων στοιχείων,
στ) η υλική αντιστοιχία μεταξύ του επιδιωκόμενου περιουσιακού οφέλους και της προκληθείσας περιουσιακής ζημίας

Υποκειμενική υπόσταση:

α) δόλος
β) σκοπός του δράστη να αποκομίσει παράνομο περιουσιακό όφελος (έγκλημα υπερχειλούς υποκειμενικής υπόστασης).

Παράσταση → η ανακοίνωση σε κάποιον μιας σκέψης ή η βεβαίωση ή ο ισχυρισμός σχετικά με ένα γεγονός
Απόκρυψη → θετική συμπεριφορά με την οποία ο παραπλανώμενος εμποδίζεται να πληροφορηθεί την αλήθεια
Παρασιώπηση η παράλειψη ανακοίνωσης αληθινών γεγονότων → όταν από τον νόμο ή τη σύμβαση ή από προηγούμενη ενέργεια του δράστη υπάρχει υποχρέωση ανακοίνωσή τους κατ' άρθρο 15 ΠΚ
Πλάνη → κάθε παράσταση στη συνείδηση του διαθέτοντας ως προς συγκεκριμένο πραγματικό περιστατικό, που δεν ανταποκρίνεται στην πραγματικότητα

ΠΗΓΗ: https://www.experian.com/blogs/global-insights/3-things-businesses-can-do-to-prevent-deepfake-fraud/

ΠΑΡΑΔΕΙΓΜΑ ΑΠΑΤΗΣ ΜΕ DEEPFAKE

Απάτη νομικών προσώπων:

- ✓ Βρετανική εταιρεία τον Μάρτιο 2019 εξαπατήθηκε κατά 250.000 ευρώ μέσω της χρήσης deepfake audio.
- ✓ Απατεώνες χρησιμοποίησαν τεχνητή νοημοσύνη για να μιμηθούν τη φωνή του Γερμανού διευθύνοντος συμβούλου της εταιρείας.
- ✓ Χρησιμοποιώντας αυτή τη φωνή, κάλεσαν έναν ανώτερο υπάλληλο και του ζήτησαν να μεταφέρει αμέσως 250.000 ευρώ στο λογαριασμό ενός υποτιθέμενου προμηθευτή ενέργειας.
- ✓ Δεδομένης της εξέχουσας θέσης του, η φωνή του ως δεδομένο θα μπορούσε να είναι δημόσια διαθέσιμη και εύκολα προσβάσιμη.
- ✓ Με τη δυνατότητα κλοπής της φωνής και της εικόνας κάποιου, η απάτη μέσω μιμήσεων έχει ενισχυθεί.
- ✓ Εντός τεσσάρων μηνών από την υπόθεση που έγινε πρωτοσέλιδο τον Μάρτιο του 2019, η εταιρεία κυβερνοασφάλειας Symantec ανέφερε ότι τρεις άλλες εταιρείες είχαν πέσει θύματα παρόμοιων παγίδων, με την Al να χρησιμοποιείται για να κλωνοποιεί φωνές και να καλεί ανώτερους οικονομικούς υπαλλήλους ζητώντας επείγουσες μεταφορές χρημάτων.

Απάτη φυσικών προσώπων:

- ✓ Έχει παρατηρηθεί το φαινόμενο, ακόμα και στην Ελλάδα, κάποιος να παριστάνει έναν επιχειρηματία ή έναν διάσημο σε κάποια θαυμάστρια και μέσω της επικοινωνίας να αποσπά τεράστια χρηματικά ποσά

ΠΗΓΗ: https://www.biometricupdate.com/202109/the-deepfake-scape-how-to-fight-fraud-in-the-digital-age

ΤΑ DEEPFAKES ΩΣ ΜΕΣΟ ΠΛΑΣΤΟΓΡΑΦΙΑΣ

Άρθρο 216 ΠΚ "1. Όποιος καταρτίζει πλαστό ή νοθεύει έγγραφο με σκοπό να παραπλανήσει με τη χρήση του άλλον σχετικά με γεγονός που μπορεί να έχει έννομες συνέπειες τιμωρείται με φυλάκιση και χρηματική ποινή…»

Αντικειμενική Υπόσταση:

Έννομο αγαθό → η αποδεικτική ακεραιότητα του «υπομνήματος», η γνησιότητα του εγγράφου

Κατάρτιση πλαστού → εξαρχής δημιουργείται έγγραφο, το οποίο προέρχεται από τρίτο πρόσωπο

Νόθευση → υλική επέμβαση σε ήδη υπάρχον έγγραφο, του οποίου μεταβάλλεται το περιεχόμενο σε ορισμένα σημεία

Υποκειμενική υπόσταση: δόλος και σκοπός του δράστη να παραπλανήσει άλλον σχετικά με γεγονός που μπορεί να έχει έννομες συνέπειες (έγκλημα υπερχειλούς υποκειμενικής υπόστασης).

Πλαστογραφία και deepfake → **κοινό συνθετικό πλαστό = fake = ψεύτικο.**

Πλαστογραφία διαβατηρίων ή ταυτοτήτων → Τα συνθετικά μέσα και οι ψηφιακά επεξεργασμένες εικόνες προσώπου αποτελούν μια νέα προσέγγιση για την πλαστογραφία εγγράφων.

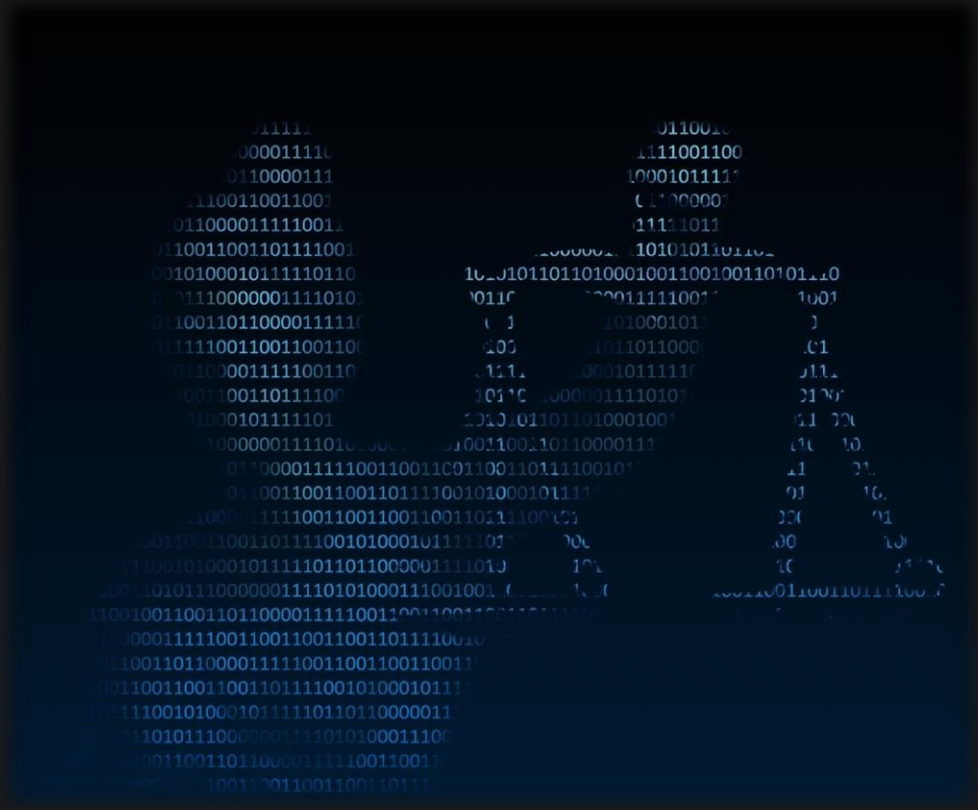
Η πλαστογραφία με έγγραφα διευκολύνει άλλα εγκλήματα, όπως η παράνομη μετανάστευση, η εμπορία ανθρώπων, η πώληση διαφόρων παράνομων αγαθών και η τρομοκρατία, καθώς οι δράστες χρησιμοποιούν συχνά πλαστές ταυτότητες για να ταξιδέψουν στις τοποθεσίες προορισμού τους.

Η τεχνολογία Deepfake ενδέχεται να ενισχύσει τον κίνδυνο προηγμένης πλαστογραφίας εγγράφων από ομάδες οργανωμένου εγκλήματος ή αντίθετα να ενισχύσει τα συστήματα αναγνώρισης πλαστών εγγράφων μέσω της αναγνώρισης προσώπου.

ΠΗΓΗ https://shuftipro.com/blog/synthetic-identity-a-new-form-of-identity-fraud-for-2019/

ΤΑ DEEPFAKES ΩΣ ΑΠΟΔΕΙΚΤΙΚΟ ΣΤΟΙΧΕΙΟ ΣΤΗ ΔΙΚΑΣΤΗΡΙΑΚΗ ΠΡΑΚΤΙΚΗ

- ✓ **Ως ψηφιακά πειστήρια νοούνται τα δεδομένα (ευρήματα) που είναι σε ψηφιακή μορφή, των οποίων ο εντοπισμός, η εξαγωγή και η ερμηνεία συμβαίνει κατά τη διαδικασία ψηφιακής σήμανσης (digital forensics).**
- ✓ **Τρόποι:** α) Κατασκευή από τον έναν διάδικο στην προσπάθεια να επικρατήσει στη δίκη ή β) Κατασκευή από τον αντίπαλο – αντίδικο
- ✓ **Ηνωμένο Βασίλειο** → η μητέρα του παιδιού εισήγαγε στη δίκη ένα παραποιημένο αρχείο ήχου για να στηρίξει τον ισχυρισμό της ότι ο πατέρας ήταν πολύ βίαιος για να του επιτραπεί η επικοινωνία με τα παιδιά τους, χρησιμοποίησε λογισμικό και διαδικτυακά σεμινάρια για να φτιάξει ένα αληθοφανές αρχείο ήχου", το οποίο ακουγόταν σαν ηχογράφηση του πατέρα που την απειλούσε σε ένα τηλεφώνημα.
- ✓ Απαιτήσεις επικύρωσης όλων των αποδεικτικών στοιχείων → Για τα βίντεο ο ρόλος του μάρτυρα είναι καθοριστικός
- ✓ Εντοπισμός των χειραγωγημένων αποδεικτικών μέσων με εργαλεία Τεχνητής Νοημοσύνης.
- ✓ Καθυστέρηση στην απονομή δικαιοσύνης και Παράταση δικαστικής διαμάχης και αύξηση του κόστους μέσω της πρόσθετης επιμέλειας για την επαλήθευση της γνησιότητας εγγράφων → Δημιουργία Καταλόγου Πραγματογνωμόνων, Διορισμός Πραγματογνώμονα, Διενέργεια Πραγματογνωμοσύνης κλπ.



ΠΗΓΗ: https://tictac.gr/online-apates/psifiaka-pistiria-ereuna-dikastikis-xrisis/

ΕΠΙΛΟΓΟΣ - ΣΥΜΠΕΡΑΣΜΑ

DEEPFAKES ΟΦΕΛΟΣ Η ΑΠΕΙΛΗ;

- Τα πάντα γύρω μας είναι τεχνολογία και κάθε νέα τεχνολογία μόνο βελτίωση επιδέχεται.
- Όλες οι τεχνολογίες δημιουργούνται με σκοπό να παρέχουν όφελος στην κοινωνία. Ισχύει για πάντα αυτό;
- Δύναμη των social media στη διάδοση εικόνων, βίντεο, ειδήσεων.
- Ευκολοπιστία ανθρώπων – Έλλειψη διασταύρωσης της πηγής της είδησης.
- Γνώση για τα Deepfakes → επαγρύπνηση για την κοινωνία πληροφορίας ή πρόκληση αναταραχών και κίνδυνοι;
- Θετικά ή αρνητικά περισσότερα για την τεχνολογία;
- Τα όπλα δεν σκοτώνουν ανθρώπους, οι άνθρωποι σκοτώνουν ανθρώπους.
- Τι ρόλο μπορεί να διαδραματίσει η Νομική Επιστήμη;
- Πρέπει να ρυθμιστεί η ίδια η τεχνολογία ή οι επιβλαβείς σκοποί της;
- Τι μέτρα, λοιπόν, πρέπει να ληφθούν για να αντιμετωπιστεί το φαινόμενο της κακόβουλης χρήσης των deepfakes;
- Επαρκεί το υφιστάμενο ποινικό οπλοστάσιο;
- Θέσπιση ειδικής νομοθεσίας ή αναδιαμόρφωση των ήδη υφιστάμενων νομικών διατάξεων;
- Άγνοια περισσότερων για τις νέες τεχνολογίες, για το πως συλλέγονται τα δεδομένα, για το ότι μπορεί να πέσουν θύματα, για το ότι η συμπεριφορά τους είναι παράνομη.

ΠΡΟΤΑΣΕΙΣ

- **Deepfake Detection** - Πρόκληση για ανάπτυξη τεχνολογιών εντοπισμού πλαστών βίντεο
- **Νέα ποινική διάταξη** με στόχο α) την ευαισθητοποίηση, β) την ενημέρωση γ) την πρόληψη δ) την προστασία.
- **Αναδιάρθρωση του ήδη υφιστάμενου ποινικού οπλοστάσιου.**
- **Διασυνοριακός Χαρακτήρας** → Συνεργασία των χωρών και των αρχών - αστυνομικών και δικαστικών - προκειμένου να δοθεί λύση στο πιο νομικό καθεστώς πρέπει να εφαρμοστεί.
- **Απαγόρευση ή Ανάπτυξη Κανόνων** ως προς τη χρήση της τεχνολογίας.
- Εκ των προτέρων **έλεγχος νομιμότητας** προτού το περιεχόμενο διανεμηθεί μεταξύ φίλων ή τεθεί στο διαδίκτυο
- **Εκστρατεία ευαισθητοποίησης**

Τρόποι προστασίας των θυμάτων

- **Διάδοση πλαστού βίντεο σε διαδικτυακή πλατφόρμα** Τι συμβαίνει με τους χρήστες των ίδιων κοινωνικών πλατφορμών σε άλλη χώρα, για τους οποίους δεν ισχύει η ίδια απαγόρευση; → Τι θα συμβεί αν σε κάποια χώρα η συμπεριφορά αυτή δεν θεωρείται αξιόποινη; → **συνεργασία των αρχών και των παρόχων σε όλες τις χώρες**
- **Υπολογιστικό νέφος** → Τι συμβαίνει όταν ο δράστης έχει αποθηκεύσει το παράνομο υλικό σε υπολογιστικό νέφος, στο οποίο έχει πρόσβαση μόνο ο ίδιος με τους προσωπικούς του κωδικούς και κανένας άλλος, με αποτέλεσμα το βίντεο να κοινοποιηθεί ξανά σε μια νέα και διαφορετική πλατφόρμα στο ίντερνετ; → **οριστική διαγραφή του προσωπικού λογαριασμού νέφους του δράστη**
- Κατά πόσο αποκαθίσταται η φήμη του θύματος όταν κάποιος χρήστης πρόλαβε να δει το παράνομο υλικό, το αποθήκευσε στην προσωπική του συσκευή ή το διέδωσε μέσω chat στα social media στους φίλους του;

ΣΑΣ ΕΥΧΑΡΙΣΤΩ!!!