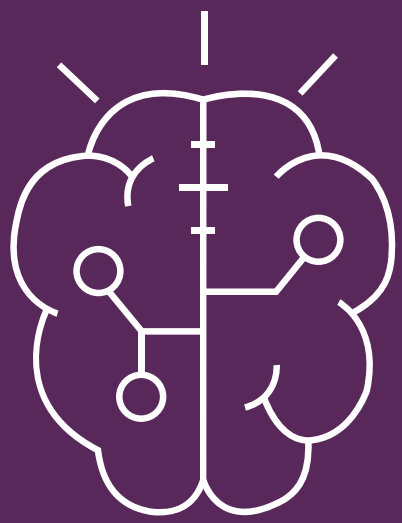
Οδηγός Εκπαιδευτικού

# Digital Well being



## Nudification & Deepfakes

Πίσω από την εικόνα: πώς να διαβάζετε ανάμεσα στις γραμμές στην εποχή των deepfakes

## Σύντομη Περιγραφή:

Οι μαθητές/τριες μαθαίνουν να αξιολογούν πιθανά deepfakes εξετάζοντας τόσο το κοινωνικό πλαίσιο και την πρόθεση του μηνύματος όσο και την αξιοπιστία της πηγής, χωρίς να βασίζονται αποκλειστικά στον εντοπισμό τεχνικών ατελειών.

[1] Deepfakes: συνθετικό ή παραποιημένο ψηφιακό περιεχόμενο (εικόνα, βίντεο ή ήχος), που δημιουργείται ή τροποποιείται με τεχνητή νοημοσύνη ώστε να παρουσιάζει πειστικά κάτι που δεν συνέβη πραγματικά.

### Απευθύνεται σε μαθητές/τριες:

- Ηλικία 13–16
- Βαθμίδα: Γυμνάσιο / Λύκειο
- Τάξεις: Α΄ Γυμνασίου έως Α΄ Λυκείου
- Επίπεδο Δυσκολίας: Μεσαίο
- Διάρκεια: 45 λεπτά (έως 60 λεπτά)

### Γνωστικά Πεδία

- Ψηφιακή ευημερία
- Κανόνες διαδικτυακής συμπεριφοράς
- Τεχνητή νοημοσύνη
- Ασφάλεια
- Ψηφιακή ταυτότητα
- Εγγραμματισμός δεδομένων και ενσυναίσθηση

### Προαπαιτούμενες Γνώσεις

- Βασική γνώση του ψηφιακού περιβάλλοντος (κοινωνικά δίκτυα, εφαρμογές, βίντεο)

## Μαθησιακοί Στόχοι & Δεξιότητες

### Στόχοι:

Στο τέλος του μαθήματος, οι μαθητές/τριες θα μπορούν να:

- Προστατεύουν την ψηφιακή τους φήμη, αξιολογώντας το κοινωνικό πλαίσιο και την πρόθεση πίσω από τα συνθετικά μέσα[2], ώστε να αναγνωρίζουν πιθανές επιβλαβείς συνέπειες.

[2] Τα συνθετικά μέσα (synthetic media) είναι ψηφιακό περιεχόμενο, όπως εικόνες, βίντεο, ηχητικά αρχεία ή κείμενα, που έχει δημιουργηθεί ή τροποποιηθεί με τη βοήθεια τεχνητής νοημοσύνης. Παραδείγματα είναι τα deepfakes, οι φωνές που μιμούνται υπαρκτά πρόσωπα, οι εικόνες που δημιουργούνται από AI και τα βίντεο στα οποία αλλάζουν τα λόγια ή οι κινήσεις κάποιου.

- Κατανοούν ότι ο κυβερνοεκφοβισμός μπορεί να περιλαμβάνει τη χρήση deepfake για ταπείνωση, πλαστοπροσωπία και κοινωνικές συγκρούσεις μέσω αλλοιωμένου βίντεο, εικόνας ή ήχου.
- Γνωρίζουν πώς να ενεργούν σωστά, χρησιμοποιώντας ενδείξεις τεχνητής νοημοσύνης και συναισθηματική αυτορρύθμιση, ώστε να αποφεύγουν τη διάδοση περιεχομένου που μπορεί να βλάψει την ασφάλεια και την αξιοπρέπεια άλλων.

### Ψηφιακές Δεξιότητες:

**Επικοινωνία και συνεργασία**

- Διαχείριση ψηφιακής ταυτότητας

**Ασφάλεια**
- Προστασία υγείας και ευημερίας

## Άλλες Δεξιότητες (Διεπιστημονικές):

- Ικανότητα ανάλυσης και σύνοψης
- Προφορική και γραπτή επικοινωνία
- Κριτική σκέψη
- Ηθική δέσμευση
- Προσαρμοστικότητα σε νέες καταστάσεις
- Ικανότητα αυτοαξιολόγησης

# Παιδαγωγική Περιγραφή και Μεθοδολογία

## Σύνοψη μαθήματος:

Το μάθημα έχει σχεδιαστεί για να βοηθήσει του/τις μαθητές/τριες να αντιμετωπίζουν με κριτική σκέψη πιθανά deepfake, δίνοντας προτεραιότητα στην ανάλυση του κοινωνικού πλαισίου και της πρόθεσης του μηνύματος και όχι στις τεχνικές ατέλειες.

Ο στόχος είναι να κατανοήσουν ότι οι επιπτώσεις των deepfakes δεν περιορίζονται στην εξαπάτηση, αλλά περιλαμβάνουν επίσης την ταπείνωση, την απώλεια εμπιστοσύνης και τις κοινωνικές συγκρούσεις. Ταυτόχρονα μαθαίνουν να αντιδρούν υπεύθυνα ακόμη και όταν δεν μπορούν να επαληθεύσουν άμεσα την αλήθεια.

Μέσα από το παιχνίδι «Fakemetric» και τη συνεργατική συζήτηση, οι μαθητές/τριες ξεφεύγουν από την απλή ανίχνευση και αναπτύσσουν:

- κριτική σκέψη
- συλλογική ευθύνη

Με αυτόν τον τρόπο, η περιέργεια για την τεχνητή νοημοσύνη μετατρέπεται σε σεβασμό και προστασία της αξιοπρέπειας των άλλων.

## Μεθοδολογία:

- **Διαδραστική παρουσίαση**
  Ξεκινάτε με εισαγωγή στα deepfake και τα συνθετικά μέσα (synthetic media) μέσω παρουσίασης, εξερευνώντας τους διαφορετικούς τύπους παραποίησης και τους κοινωνικούς κινδύνους που ενδέχεται να προκαλέσουν.
- **Παιχνίδι «Fakemetric»**
  Πραγματοποιείτε μια συνεργατική δραστηριότητα στην οποία οι μαθητές/τριες κινούνται σε «γκρίζες ζώνες» αβεβαιότητας. Περιγράφουν φανταστικά βίντεο και αξιολογούν την πιθανότητα να είναι deepfake, ενισχύοντας την ικανότητά τους να διαχειρίζονται την αβεβαιότητα.
- **Αναστοχαστική δέσμευση**
  Ολοκληρώνετε με ομαδική συζήτηση για τις πραγματικές συνέπειες της διάδοσης αλλοιωμένου περιεχομένου και καταλήγετε σε προσωπική δέσμευση για υπεύθυνη ψηφιακή συμπεριφορά.

# Οδηγίες

## Προετοιμασία

(π.χ. εγκατάσταση λογισμικού, εκτύπωση υλικού):

- Προετοιμάστε το απαραίτητο υλικό για τη δραστηριότητα.
- Έχετε έτοιμο το εκτυπώσιμο έγγραφο με τις συμβουλές για τον εντοπισμό deepfake και τα στοιχεία του παιχνιδιού, ένα αντίτυπο για κάθε ομάδα.
- Έχετε έτοιμο το εκτυπώσιμο έγγραφο με το υποστηρικτικό υλικό της παρουσίασης, ένα αντίτυπο για κάθε εκπαιδευτικό.
- Έχετε ανοιχτή την παρουσίαση σε μια συσκευή προβολής.
- Μελετήστε την παρουσίαση πριν ξεκινήσετε τη δραστηριότητα.

## Υλικό

**Εξοπλισμός (Hardware):**

- Τάμπλετ ή φορητός υπολογιστής (μόνο για τον/την εκπαιδευτικό)

**Λογισμικό:**

- Ψηφιακός πίνακας (προαιρετικά)

**Πρόσθετα υλικά:**

- Ψαλίδια

**Εξωτερικοί σύνδεσμοι & υλικά:**

- Παρουσίαση Genially : https://view.genially.com/6a996bf39bb5591115e39aa5
- Ψηφιακός πίνακας (ενδεικτικά): https://classroomscreen.com/

# Περιεχόμενο (Βήματα και ενότητες)

## Θεωρητικό υπόβαθρο για εκπαιδευτικούς

Σήμερα, η εκπαίδευση γύρω από τα deepfake δεν μπορεί να περιορίζεται στον εντοπισμό τεχνικών ατελειών, καθώς το κατασκευασμένο και αλλοιωμένο (συνθετικό) περιεχόμενο γίνεται όλο και πιο πειστικό. Αυτή η δραστηριότητα προτείνει μια αλλαγή στρατηγικής: να μάθουν οι μαθητές/τριες να εξετάζουν το πλαίσιο και όχι μόνο την εικόνα.

Στόχος είναι να μάθουν να είναι επιφυλακτικοί όταν ένα βίντεο προσπαθεί να προκαλέσει έντονη συναισθηματική αντίδραση ή όταν η πηγή δεν είναι αξιόπιστη.

Το παιχνίδι «Fakemetric» είναι κεντρικό σε αυτή τη διαδικασία, γιατί τους φέρνει αντιμέτωπους με καταστάσεις αβεβαιότητας. Στο διαδίκτυο, τα πράγματα είναι σπάνια απολύτως βέβαια και η ικανότητα να διαχειρίζονται αυτή την αβεβαιότητα είναι που οδηγεί στην ανάπτυξη πραγματικής κριτικής σκέψης.

Μέσα από τη συζήτηση σε ομάδες, οι μαθητές/τριες μετατρέπονται από παθητικοί καταναλωτές σε ενεργούς αναλυτές που αμφισβητούν την πρόθεση αυτού που βλέπουν.

Τέλος, η δραστηριότητα συνδέει την τεχνολογία με την ηθική και τη συνύπαρξη. Εστιάζοντας στους κινδύνους των deepfake στο σχολικό περιβάλλον, ενισχύεται η ατομική ευθύνη: η δέσμευση να μην διαδίδουν περιεχόμενο που μπορεί να βλάψει άλλους. Έτσι, η περιέργεια για την τεχνητή νοημοσύνη μετατρέπεται σε εργαλείο συλλογικής προστασίας και ψηφιακού σεβασμού.

**Σημείωση προς τον/την εκπαιδευτικό: Ασφάλεια και βασικοί κανόνες**
Υπενθυμίστε στους μαθητές/τριες:

- Να μην αναφέρουν πραγματικά άτομα, σχολεία ή λογαριασμούς.
- Να μην αναφέρουν εφαρμογές ή εργαλεία δημιουργίας και διάδοσης deepfake.
- Να μην χρησιμοποιούν παραδείγματα με αναγνωρίσιμους ανηλίκους ή προσωπικό του σχολείου.
- Να αποφεύγεται κάθε σεξουαλικοποιημένο περιεχόμενο που αφορά ανηλίκους.

Όλα τα παραδείγματα πρέπει να είναι κατάλληλα για το σχολικό περιβάλλον και επιλεγμένα από τον/την εκπαιδευτικό.

Η κοινοποίηση προσωπικών εμπειριών είναι προαιρετική. Αν κάτι προκαλέσει αναστάτωση ή σχετίζεται με πραγματικά γεγονότα, οι μαθητές/τριες μπορούν να μιλήσουν ιδιωτικά με τον/την εκπαιδευτικό, τον/τη διευθυντή/ντρια ή τον/την υπεύθυνο/η προστασίας του σχολείου. Κανείς δεν τιμωρείται όταν ζητά βοήθεια. Τονίστε ότι η αναζήτηση βοήθειας δεν θα οδηγήσει σε τιμωρία και ότι στόχος του μαθήματος είναι η ανάπτυξη ασφαλούς κρίσης.

**Χρήσιμες γραμμές βοήθειας**
- Ελληνικό Κέντρο Ασφαλούς Διαδικτύου / SaferInternet4Kids: Γραμμή Βοήθειας Help-line.gr: 210 6007686
- «Το Χαμόγελο του Παιδιού»: Εθνική Γραμμή για τα Παιδιά SOS 1056 και Ευρωπαϊκή Γραμμή Υποστήριξης Παιδιών 116111
- Διεύθυνση Δίωξης Ηλεκτρονικού Εγκλήματος: 11188

**Υποστηρικτικές υπηρεσίες και εργαλεία**
- Tozi: δωρεάν εφαρμογή του Ιδρύματος Vodafone για την ψηφιακή ευημερία και την αντιμετώπιση του διαδικτυακού εκφοβισμού. Απευθύνεται σε παιδιά ηλικίας 11+ και προσφέρει υποστήριξη και εργαλεία για μια πιο ασφαλή εμπειρία στον ψηφιακό κόσμο.
  https://www.to-zi.com/welcome-landing-el.html

## Υλοποίηση

### Εισαγωγή (10 λεπτά)

Ξεκινήστε το μάθημα προβάλλοντας τη διαδραστική παρουσίαση, στην οποία οι μαθητές/τριες θα δουν βίντεο deepfake.

Αν δεν υπάρχει προβολέας, οργανώστε την τάξη σε ομάδες και δώστε σε κάθε ομάδα έναν υπολογιστή, ανοίγοντας εκεί την παρουσίαση και ζητώντας τους να σας ακολουθούν.

Σε κάθε ερώτηση:

- σταματήστε

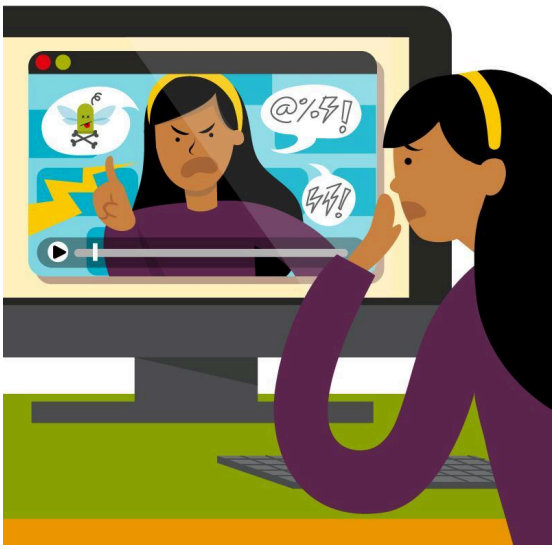
- ζητήστε απόψεις
- δώστε τον λόγο σε παιδιά που θέλουν να μιλήσουν

Οι ερωτήσεις ακολουθούν μια κοινή λογική πορεία, ώστε οι μαθητές/τριες να οδηγηθούν στο συμπέρασμα ότι: τα deepfake γίνονται όλο και πιο εξελιγμένα τεχνολογικά και ότι ο τρόπος για να καταλάβουμε αν ένα βίντεο είναι αληθινό δεν είναι να εντοπίζουμε μόνο τεχνικές ατέλειες, αλλά να κατανοούμε το πλαίσιο και τι κρύβεται πίσω από αυτό.

Το περιεχόμενο της παρουσίασης περιλαμβάνει τα εξής θέματα:

- Τι γνωρίζουμε για τα deepfake;
- Ποιοι τύποι deepfake υπάρχουν;
- Συνήθως βλέπουμε δημόσια πρόσωπα ως παραδείγματα, αλλά μπορούν τα deepfake να επηρεάσουν και άτομα στο σχολείο ή στην παρέα μας; Αν ναι, με ποιον τρόπο;
- Ποια σημάδια μπορούν να μας κάνουν να αμφισβητήσουμε την αυθεντικότητα;
- Τι είναι τα υδατογραφήματα και οι ενδείξεις τεχνητής νοημοσύνης;
- Γιατί αυτά τα οπτικά στοιχεία και οι τεχνικές ατέλειες είναι χρήσιμα αλλά όχι αρκετά;
- Όταν τα οπτικά στοιχεία δεν επαρκούν, με ποιον άλλο τρόπο μπορούμε να εξετάσουμε την αξιοπιστία ενός πιθανού deepfake;
- Τι πρέπει να κάνουμε όταν δεν είμαστε σίγουροι αν ένα βίντεο είναι ψεύτικο;

Για τη συζήτηση κάθε θέματος, μπορείτε να χρησιμοποιήσετε το υποστηρικτικό υλικό της παρουσίασης.




## Το παιχνίδι Fakemetric (30 λεπτά)

Στη συνέχεια, οργανώστε ομάδες εργασίας αποτελούμενες από 4–6 μαθητές/τριες και μοιράστε σε κάθε ομάδα το υλικό και τα εκτυπωμένα φύλλα με τις συμβουλές για τον εντοπισμό deepfake και τα στοιχεία του παιχνιδιού.

Εξηγήστε ότι στο πρώτο Φύλλο θα βρουν μια λίστα συμβουλών για να εντοπίσουν αν ένα βίντεο μπορεί να είναι deepfake. Οι συμβουλές αυτές χωρίζονται σε τρεις κατηγορίες:

- **Πράγματα που παρατηρώ αμέσως:**
  Αυτή η κατηγορία επικεντρώνεται σε εξωτερικά στοιχεία του βίντεο. Είναι γρήγορες ενδείξεις σχετικά με το ποιος δημοσιεύει το περιεχόμενο και πώς το παρουσιάζει, που ενεργοποιούν την πρώτη μας υποψία.
- **Πράγματα που μου φαίνονται ύποπτα:**
  Εδώ εξετάζουμε το ίδιο το περιεχόμενο του βίντεο. Πρόκειται για λεπτές λεπτομέρειες στην εικόνα, στον ήχο ή στη συμπεριφορά που δημιουργούν την αίσθηση ότι «κάτι δεν ταιριάζει».
- **Πράγματα που πρέπει να ελέγξω:**
  Αυτή η ενότητα περιγράφει τα βήματα διερεύνησης: τη μετάβαση από την υποψία στην επιβεβαίωση, μέσω εργαλείων και σύγκρισης με αξιόπιστες πηγές.

Στη συνέχεια, εξηγήστε ότι θα εξασκηθούν στον εντοπισμό deepfake μέσα από ένα παιχνίδι που ονομάζεται «Fakemetric».

Ζητήστε από τους/τις μαθητές/τριες να κόψουν όλα τα στοιχεία του παιχνιδιού:

- ταμπλό
- δείκτη
- έντεκα κάρτες

Αφού τα ετοιμάσουν, τοποθετούν τον δείκτη στο ταμπλό, δείχνοντας σε οποιοδήποτε ποσοστό του ταμπλό. Στη συνέχεια, παίρνουν τις κάρτες, τις ανακατεύουν και τις τοποθετούν ανάποδα (με την όψη προς τα κάτω) δίπλα στο ταμπλό.




Εξηγήστε τους κανόνες του παιχνιδιού μαζί με παραδείγματα:

- Στην περίμετρο του ταμπλό υπάρχουν ποσοστά, τα οποία δείχνουν πόσο προσεκτικά πρέπει να εξετάζεται ένα φανταστικό βίντεο ως προς το αν είναι deepfake. Περιλαμβάνει επίσης υπενθυμίσεις των κανόνων του παιχνιδιού.
- Κάθε κάρτα έχει μια τιμή από 0% έως 100%, που αντιστοιχεί στις τιμές του ταμπλό.
- Το παιχνίδι είναι συνεργατικό. Κάθε ομάδα συνεργάζεται για να συγκεντρώσει όσο το δυνατόν περισσότερους πόντους.

- Σε κάθε γύρο, ένας μαθητής/μια μαθήτρια παίρνει μια κάρτα χωρίς να τη δείξει στους υπόλοιπους και τη διαβάζει μόνος/μόνη του.

Παράδειγμα:

Ένας μαθητής/μια μαθήτρια τραβάει κρυφά την κάρτα 100%.

- Ο στόχος του είναι να βοηθήσει την ομάδα να τοποθετήσει τον δείκτη όσο πιο κοντά γίνεται στην τιμή της κάρτας. Για να το πετύχει, περιγράφει ένα φανταστικό βίντεο που αντιστοιχεί σε αυτό το επίπεδο «ύποπτου» περιεχομένου.

Παράδειγμα περιγραφής:

Ένας πολύ γνωστός ποδοσφαιριστής εμφανίζεται ντυμένος μονόκερος και τραγουδά ένα παραδοσιακό τραγούδι. Το στόμα κινείται ελάχιστα και τα μάτια δεν ανοιγοκλείνουν. Ο λογαριασμός που ανέβασε το βίντεο λέγεται «Only-AI-fakes».

- Μετά την περιγραφή, ο μαθητής/η μαθήτρια δεν επιτρέπεται να μιλήσει ξανά.
- Η ομάδα συζητά και τοποθετεί τον δείκτη στο ποσοστό που θεωρεί πιο κατάλληλο.
- Αν η ομάδα πετύχει ακριβώς το ποσοστό της κάρτας, κερδίζει 2 πόντους.
- Αν βρεθεί σε διπλανή τιμή, κερδίζει 1 πόντο.

Παράδειγμα:

Αν τοποθετήσουν τον δείκτη στο 100%, παίρνουν 2 πόντους. Αν τον τοποθετήσουν στο 90%, παίρνουν 1 πόντο.

- Το παιχνίδι συνεχίζεται με νέο γύρο και νέο μαθητή/μαθήτρια.
- Όταν όλοι έχουν παίξει έναν γύρο, το παιχνίδι ολοκληρώνεται.
- Το τελικό σκορ είναι το άθροισμα όλων των γύρων.

Πιο σύνθετο παράδειγμα:

Ένας μαθητής τραβάει την κάρτα 40% και περιγράφει ένα βίντεο όπου μια γνωστή τραγουδίστρια ανακοινώνει ότι θα δώσει δωρεάν συναυλία σε μία ώρα. Υπάρχουν πολλές αναρτήσεις, αλλά όχι επιβεβαίωση από επίσημους λογαριασμούς.

Η ομάδα συζητά: θεωρεί απίθανο (και ριψοκίνδυνο) μια τόσο διάσημη τραγουδίστρια να δώσει αυθόρμητη δωρεάν συναυλία, όμως φαίνεται ότι πολλές πηγές το αναπαράγουν, χωρίς επιβεβαίωση από τους επίσημους λογαριασμούς της. Τελικά τοποθετεί τον δείκτη στο 50%. Ο μαθητής αποκαλύπτει την κάρτα και η ομάδα κερδίζει 1 πόντο.

Υπενθυμίστε ότι μπορούν να χρησιμοποιούν τις συμβουλές από το εκτυπώσιμο υλικό.

Αφήστε τους τον υπόλοιπο χρόνο του μαθήματος για να παίξουν έναν ή δύο γύρους ακόμα. Αν υπάρχει χρόνος, μπορείτε να αναδιοργανώσετε τις ομάδες.

# Συμπεράσματα και μελλοντικές προοπτικές:

## Συμπέρασμα (10 λεπτά)

Λίγο πριν το τέλος του μαθήματος, ζητήστε από τους μαθητές/τριες να σταματήσουν το παιχνίδι και να μοιραστούν τη βαθμολογία τους με τις κοντινές ομάδες.

Εξηγήστε τους ότι, κατά τη διάρκεια του παιχνιδιού, μπόρεσαν να διαπιστώσουν ότι είναι δύσκολο να ξεχωρίσουμε με βεβαιότητα αν ένα βίντεο είναι ψεύτικο ή όχι.

Στη συνέχεια, ζητήστε τους να προβληματιστούν πάνω στα εξής θέματα:

- Ποιοι κίνδυνοι υπάρχουν όταν διαμοιραζόμαστε ένα deepfake;
- Είναι όλα τα deepfake το ίδιο επιβλαβή; Υπάρχουν κάποια που είναι πιο «αθώα» από άλλα; Ζητήστε τους να δώσουν παραδείγματα.

Τέλος, ολοκληρώστε το μάθημα με μία σημαντική δέσμευση:

Ανεξάρτητα από το αν ένα βίντεο είναι αληθινό ή συνθετικό, σε περίπτωση που να βλάψει την αξιοπρέπεια, την ασφάλεια ή την αξιοπιστία κάποιου ατόμου, δεν το διαδίδουμε. Αντίθετα, σταματάμε πριν αντιδράσουμε και ενημερώνουμε έναν έμπιστο ενήλικα και, αν χρειάζεται, το άτομο που επηρεάζεται.

**Πληροφορίες για εκπαιδευτικούς**
Η κοινοποίηση ενός deepfake μπορεί να προκαλέσει πραγματική και συχνά μη αναστρέψιμη βλάβη.

Ο βασικός κίνδυνος είναι η πλαστοπροσωπία, η οποία χρησιμοποιείται για:

- κατασκευή ψευδών αποδείξεων
- καταστροφή της δημόσιας εικόνας κάποιου
- διάπραξη απάτης

Με τη διάδοση τέτοιου περιεχομένου, γινόμαστε συνυπεύθυνοι για την επικύρωση ενός ψέματος που μπορεί να οδηγήσει σε σοβαρές προσωπικές ή ακόμα και νομικές συνέπειες για το άτομο που επηρεάζεται.

Μόλις το περιεχόμενο κοινοποιηθεί, χάνουμε τον έλεγχο πάνω σε αυτό και η ζημιά στην ιδιωτικότητα ή την ψυχική υγεία του θύματος πολλαπλασιάζεται και παραμένει.

Ωστόσο, δεν έχουν όλα τα deepfake τον ίδιο αντίκτυπο.

Από τη μία πλευρά, υπάρχουν δημιουργικά ή «αθώα» παραδείγματα, όπως:

- φίλτρα που τοποθετούν το πρόσωπό μας σε χαρακτήρες ταινιών
- σατιρικά βίντεο όπου είναι ξεκάθαρο ότι πρόκειται για παρωδία (π.χ. ιστορικό πρόσωπο που τραγουδά ένα σύγχρονο τραγούδι)

Αυτά θεωρούνται πιο ασφαλή, γιατί δεν στοχεύουν στην εξαπάτηση.

Από την άλλη πλευρά, υπάρχουν επιβλαβή deepfake, σχεδιασμένα για να:

- βλάψουν
- χειραγωγήσουν

Παραδείγματα περιλαμβάνουν:

- βίντεο που αποδίδουν ψευδείς δηλώσεις σε δημόσια πρόσωπα για δημιουργία σύγκρουσης
- υλικό που γελοιοποιεί ή παρενοχλεί κάποιο άτομο στο σχολικό περιβάλλον

Το βασικό μήνυμα που πρέπει να μεταδοθεί είναι το εξής:

Αν ένα βίντεο χρησιμοποιεί την εικόνα κάποιου για να τον παρουσιάσει να κάνει κάτι ταπεινωτικό ή επικίνδυνο, τότε δεν πρόκειται για αστείο αλλά για εργαλείο πρόκλησης βλάβης και δεν πρέπει ποτέ να το διαδίδουμε.

## Επέκταση ή διαφοροποίηση / Περαιτέρω ιδέες

Αν υπάρχει περισσότερος χρόνος διαθέσιμος ή αν οι μαθητές/τριες προχωρούν πιο γρήγορα από το αναμενόμενο:

- Πριν την έναρξη, προσθέστε περισσότερα παραδείγματα βίντεο deepfake για προβολή στην αρχική παρουσίαση, ώστε να εμπλουτιστεί η εισαγωγή.
- Παίξτε περισσότερους γύρους του παιχνιδιού «Fakemetric».

## Λιγότερος χρόνος;

Αν υπάρχει λιγότερος χρόνος διαθέσιμος:

- Μπορείτε να παίξετε λιγότερους γύρους του «Fakemetric».
- Μπορείτε να δημιουργήσετε μικρότερες ομάδες, μειώνοντας έτσι τη διάρκεια κάθε παιχνιδιού.

## Αξιολόγηση

Για να διασφαλίσετε ότι οι μαθητές/τριες έχουν επιτύχει τους στόχους της δραστηριότητας, προτείνεται να τους παρατηρείτε κατά τη διάρκεια του μαθήματος και να εξετάζετε αν:

- Συμμετέχουν ενεργά στις συζητήσεις και στις ερωτήσεις της αρχικής παρουσίασης.
- Περιγράφουν βίντεο deepfake σύμφωνα με τις κάρτες που έχουν επιλέξει και χρησιμοποιούν τις συμβουλές από το εκτυπώσιμο υλικό.

Επίσης, είναι σημαντικό να αξιολογήσετε:

- αν οι μαθητές/τριες απόλαυσαν τη δραστηριότητα
- αν ο χρόνος και το παρεχόμενο υλικό ήταν κατάλληλα για την υλοποίησή της

# Παράρτημα 1 - Συμβουλές για τον εντοπισμό deepfakes

## 1 Πράγματα που παρατηρώ αμέσως

- Έχει δημοσιευτεί από έναν νέο, ανώνυμο ή ασυνήθιστο λογαριασμό.
- Το βίντεο παρουσιάζεται ως αναδημοσίευση είδησης ή ρεπορτάζ, αλλά δεν μπορεί να βρεθεί η αρχική πηγή.
- Δεν υπάρχει σαφής ημερομηνία, τοποθεσία, γεγονός ή πηγή που να συνδέεται με το βίντεο.
- Ο τίτλος ή η λεζάντα είναι υπερβολικά συναισθηματική ή επείγουσα (π.χ. «μοιράσου το τώρα»).
- Ο λογαριασμός που το μοιράζεται έχει ιστορικό από αστεία, μοντάζ ή εντυπωσιοθηρικές δημοσιεύσεις.
- Η πλατφόρμα εμφανίζει ένδειξη τεχνητής νοημοσύνης ή σήμανση συνθετικού περιεχομένου (αυτό είναι χρήσιμο, αλλά δεν εμφανίζεται σε όλο το περιεχόμενο, είτε αληθινό είτε ψεύτικο).

## 2 Πράγματα που μου φαίνονται ύποπτα

- Η φωνή ακούγεται κάπως επίπεδη, ρομποτική ή με περίεργο ρυθμό.
- Η κίνηση του προσώπου, του στόματος ή του σώματος φαίνεται αφύσικη, ειδικά όταν το άτομο κινείται ή μιλά γρήγορα.
- Ο φωτισμός ή τα περιγράμματα γύρω από το πρόσωπο και τα μαλλιά δεν ταιριάζουν με την υπόλοιπη εικόνα.
- Το άτομο φαίνεται να λέει ή να κάνει κάτι που δεν συνάδει με τον χαρακτήρα του ή είναι υπερβολικό.
- Το βίντεο φαίνεται σχεδιασμένο για να προκαλέσει άμεση αντίδραση, όπως σοκ, θυμό ή ντροπή.
- Το αποτέλεσμα φαίνεται πολύ «τέλειο», αλλά ο συνδυασμός εικόνας, ήχου και πλαισίου δεν δένει πλήρως.

## 3 Πράγματα που πρέπει να ελέγξω

- Αναζητήστε την αρχική πηγή: πού εμφανίστηκε πρώτα και ποιος το δημοσίευσε;
- Κάντε αντίστροφη αναζήτηση εικόνας από στιγμιότυπο ή καρέ του βίντεο, όταν είναι δυνατόν.
- Ελέγξτε αν το ίδιο βίντεο εμφανίζεται αλλού με διαφορετικές πληροφορίες.
- Εξετάστε αν η είδηση εμφανίζεται σε αξιόπιστα μέσα ενημέρωσης ή επίσημα κανάλια και όχι μόνο σε λογαριασμούς που αναπαράγουν περιεχόμενο.
- Δείτε το ιστορικό του λογαριασμού: είναι αξιόπιστος ή δημοσιεύει κυρίως φήμες και περιεχόμενο τεχνητής νοημοσύνης;

## Να θυμάστε

Πριν μοιραστείτε ή αντιδράσετε, προσπαθήστε να επαληθεύσετε την πληροφορία μέσω μιας δεύτερης αξιόπιστης πηγής. Αν το βίντεο μπορεί να βλάψει την ασφάλεια ή τη φήμη κάποιου, η ασφαλέστερη επιλογή είναι να μην το διαδώσετε και να ενημερώσετε έναν έμπιστο ενήλικα.

# Παράρτημα 2 - Υποστηρικτικό υλικό για την παρουσίαση

## Τι είναι τα deepfake;

Όταν μιλάμε για deepfake, αναφερόμαστε σε βίντεο, εικόνες ή ηχητικά αρχεία που έχουν δημιουργηθεί ή τροποποιηθεί με τη χρήση τεχνητής νοημοσύνης, ώστε να παρουσιάζουν ανθρώπους να λένε ή να κάνουν πράγματα που δεν έχουν συμβεί ποτέ. Δεν πρόκειται για απλές «διορθώσεις» ή διασκεδαστικά φίλτρα, αλλά για συνθετικές δημιουργίες που χρησιμοποιούν προηγμένους αλγορίθμους για να μιμηθούν με μεγάλη ακρίβεια την εμφάνιση και τη φωνή ενός πραγματικού ανθρώπου. Στην καθημερινότητά μας, συναντάμε τέτοιο περιεχόμενο κυρίως στα κοινωνικά δίκτυα, άλλοτε για ψυχαγωγία και άλλοτε με πιο επιβλαβείς σκοπούς.

Ο όρος προέρχεται από τον συνδυασμό της «βαθιάς μάθησης» και της λέξης «ψεύτικο». Το βασικό χαρακτηριστικό δεν είναι μόνο ότι το αποτέλεσμα φαίνεται αληθοφανές, αλλά ότι έχει «εκπαιδευτεί» από ένα σύστημα που έχει μάθει να αναπαράγει κάθε λεπτομέρεια της εμφάνισης ή της φωνής ενός προσώπου με σκοπό να μας παραπλανήσει.

## Ποιοι τύποι deepfake υπάρχουν;

Μπορούμε να διακρίνουμε τρεις βασικές κατηγορίες που συναντούν συχνά οι μαθητές/τριες:

- Βίντεο: γίνεται αντικατάσταση προσώπου ή συγχρονισμός χειλιών με ψεύτικο λόγο.
- Στατικές εικόνες: φωτογραφίες που φαίνονται αληθινές αλλά έχουν δημιουργηθεί εξ ολοκλήρου από τεχνητή νοημοσύνη.
- Φωνή ή ήχος: αντιγράφεται ο τόνος, ο ρυθμός και ο τρόπος ομιλίας ενός ατόμου ώστε να δημιουργηθεί ψεύτικο ηχητικό μήνυμα.

Είναι σημαντικό να τονιστεί ότι αυτές οι μορφές συχνά συνδυάζονται. Για παράδειγμα, ένα βίντεο γίνεται πιο πειστικό όταν συνοδεύεται και από κλωνοποιημένη φωνή.

## Μπορούν τα deepfake να επηρεάσουν μαθητές/τριες;

Παρόλο που τα μέσα ενημέρωσης παρουσιάζουν συχνά παραδείγματα με πολιτικούς ή διάσημους, τα εργαλεία δημιουργίας deepfake είναι πλέον εύκολα προσβάσιμα. Είναι σημαντικό να συζητήσουμε ανοιχτά ότι αυτή η τεχνολογία μπορεί να χρησιμοποιηθεί για «αστεία» που στην πραγματικότητα αποτελούν μορφές παρενόχλησης, όπως:

- βίντεο που γελοιοποιούν κάποιον στην τάξη
- ψεύτικες εικόνες που μπορούν να βλάψουν σοβαρά τη φήμη και την ψυχική υγεία ενός ατόμου

Ως εκπαιδευτικοί, είναι σημαντικό να υπενθυμίζουμε ότι πίσω από κάθε οθόνη υπάρχει ένας πραγματικός άνθρωπος. Ανεξάρτητα από το αν ένα περιεχόμενο είναι αληθινό ή ψεύτικο, αν η διάδοσή του μπορεί να ταπεινώσει ή να πληγώσει κάποιον, υπάρχει η ηθική υποχρέωση να μη συμμετέχουμε στη διάδοσή του.

Στόχος είναι να καλλιεργήσουμε μια κουλτούρα φροντίδας, όπου οι μαθητές/τριες αντιλαμβάνονται ότι η χρήση ψηφιακών εργαλείων συνδέεται άμεσα με τον σεβασμό και την προστασία των άλλων.

# Παράρτημα 2 - Υποστηρικτικό υλικό για την παρουσίαση

## Ποια σημάδια μπορούν να μας κάνουν να αμφισβητήσουμε την αυθεντικότητα;

Όπως σε ένα «μαγικό κόλπο», αν παρατηρήσουμε προσεκτικά, μπορούμε μερικές φορές να εντοπίσουμε τα σημάδια της τεχνητής παρέμβασης. Τα πιο συνηθισμένα προβλήματα εντοπίζονται σε:

- Βλέμμα και μάτια: Παρατηρήστε αν το ανοιγοκλείσιμο των ματιών είναι αφύσικο ή σπάνιο. Ελέγξτε αν οι αντανακλάσεις στα μάτια ταιριάζουν με τον φωτισμό του χώρου.
- Στόμα και συγχρονισμός: Τα χείλη μπορεί να μην συγχρονίζονται σωστά με τον ήχο. Τα δόντια μπορεί να φαίνονται θολά και η γλώσσα να κινείται αφύσικα.
- Όρια και υφές: Παρατηρήστε την ένωση προσώπου με μαλλιά ή λαιμό. Μπορεί να εμφανίζονται θολώματα ή μικρές «διαταραχές». Επίσης, το δέρμα μπορεί να φαίνεται υπερβολικά τέλειο.
- Φυσική και περιβάλλον: Ελέγξτε αν οι σκιές και ο φωτισμός είναι φυσιολογικοί. Αξεσουάρ όπως γυαλιά ή σκουλαρίκια μπορεί να παραμορφώνονται κατά την κίνηση.

Είναι σημαντικό να επισημανθεί ότι αυτά τα λάθη μειώνονται καθώς η τεχνολογία εξελίσσεται. Η απουσία λαθών δεν σημαίνει ότι το βίντεο είναι αληθινό. Η παρατήρηση είναι απλώς το πρώτο βήμα. Το σημαντικότερο είναι η κατανόηση του πλαισίου.

## Τι είναι τα υδατογραφήματα και οι ενδείξεις τεχνητής νοημοσύνης;

Τα υδατογραφήματα και οι ετικέτες είναι εργαλεία διαφάνειας που χρησιμοποιούνται για να δείξουν ότι ένα περιεχόμενο είναι κατασκευασμένο/αλλοιωμένο (συνθετικό). Ορισμένες πλατφόρμες εντοπίζουν αυτόματα περιεχόμενο που έχει δημιουργηθεί με τεχνητή νοημοσύνη ή ενσωματώνουν πληροφορίες μέσα στο αρχείο. Ωστόσο:

- οι ενδείξεις μπορούν να αφαιρεθούν
- δεν χρησιμοποιούνται πάντα από όλα τα εργαλεία

Επομένως, η απουσία ένδειξης δεν σημαίνει ότι το περιεχόμενο είναι αληθινό. Οι ενδείξεις βοηθούν, αλλά δεν αντικαθιστούν την κριτική σκέψη και την ανάλυση του πλαισίου.

# Παράρτημα 2 - Υποστηρικτικό υλικό για την παρουσίαση

## Γιατί τα τεχνικά σημάδια δεν είναι αρκετά;

Η τεχνολογία εξελίσσεται ταχύτερα από την ικανότητά μας να εντοπίζουμε λάθη. Αν βασιστούμε μόνο στον εντοπισμό «σφαλμάτων», οι μαθητές/τριες θα μείνουν απροστάτευτοι όταν αυτά εξαφανιστούν.

Επιπλέον, υπάρχει το λεγόμενο «μέρισμα του ψεύτη» (liar's dividend): αν συνηθίσουμε να θεωρούμε ψεύτικο ό,τι φαίνεται περίεργο, κάποιος μπορεί να ισχυριστεί ότι ένα πραγματικό βίντεο είναι ψεύτικο για να αποφύγει τις ευθύνες του. Για αυτό, είναι σημαντικό να εστιάζουμε στο κοινωνικό πλαίσιο και όχι μόνο στην εικόνα.

## Πώς αλλιώς μπορούμε να ελέγξουμε την αξιοπιστία;

Αντί να κοιτάμε μόνο την εικόνα, εστιάζουμε σε στοιχεία όπως:

- Συναισθηματική πρόθεση: Αν ένα βίντεο προκαλεί έντονη αντίδραση ή επείγουσα ανάγκη για κοινοποίηση, μπορεί να αποτελεί ένδειξη χειραγώγησης.
- Συνέπεια: Είναι λογικό αυτό που λέγεται ή γίνεται; Ταιριάζει με το πρόσωπο και το περιβάλλον;
- Προέλευση: Υπάρχει αξιόπιστη αρχική πηγή; Υπάρχουν επιβεβαιώσεις από έγκυρα μέσα;

## Τι πρέπει να κάνουμε όταν δεν είμαστε σίγουροι;

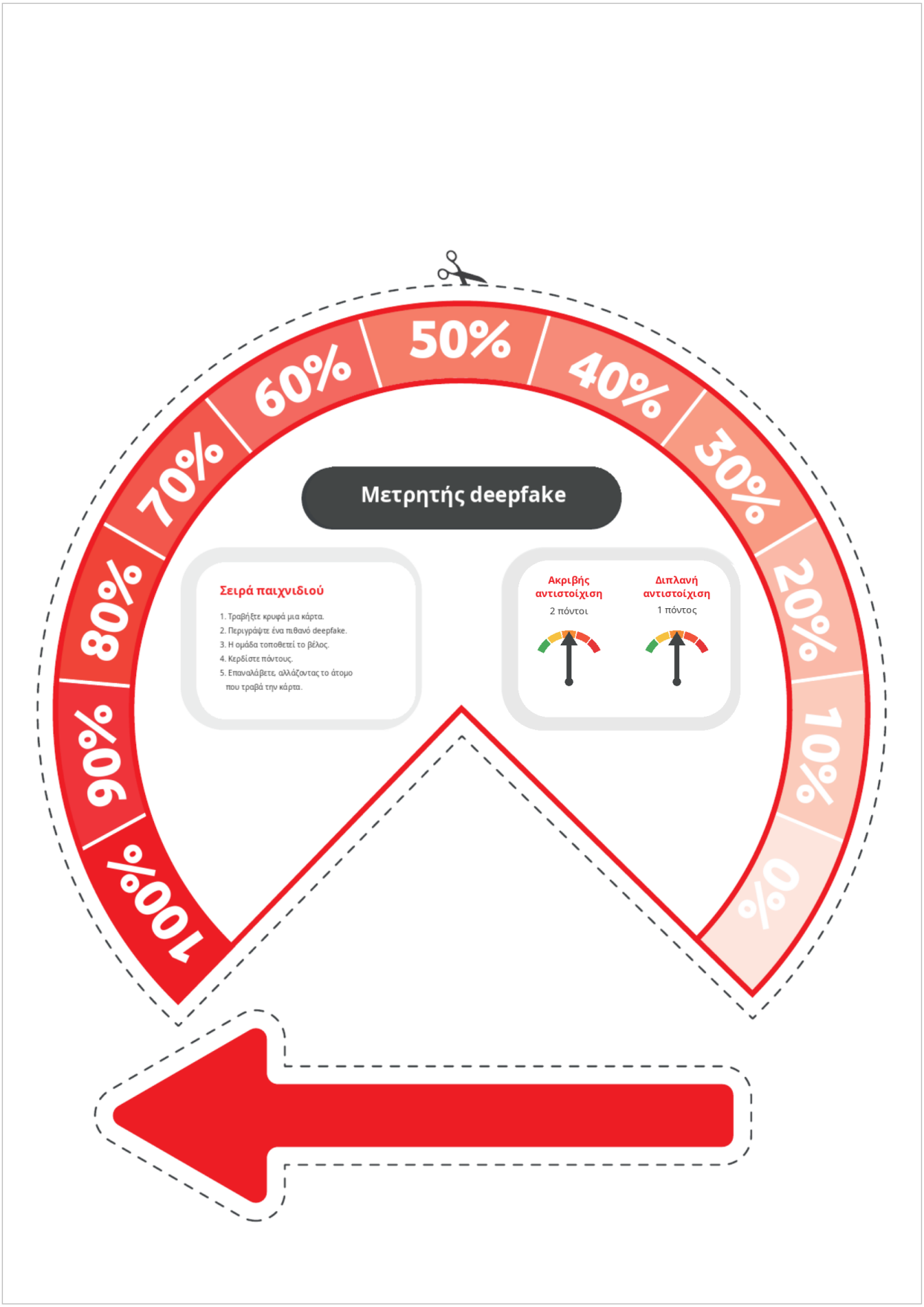
Το πιο σημαντικό βήμα είναι: Σταματάμε. Σε ένα περιβάλλον όπου όλα κινούνται γρήγορα, το να μην κοινοποιήσουμε κάτι είναι πράξη υπευθυνότητας. Δεν είναι απαραίτητο να γνωρίζουμε αμέσως αν κάτι είναι αληθινό ή ψεύτικο.

Στη συνέχεια:

- αναζητούμε την αρχική πηγή
- ελέγχουμε αν υπάρχει σχετική κάλυψη από αξιόπιστα μέσα
- χρησιμοποιούμε εργαλεία επαλήθευσης
- απευθυνόμαστε σε έναν έμπιστο ενήλικα

Η αναζήτηση βοήθειας δείχνει ψηφιακή ωριμότητα και υπευθυνότητα.

# Παράρτημα 3 - Στοιχεία του παιχνιδιού



## Παράρτημα 3 - Στοιχεία του παιχνιδιού

