

Ετήσιο Ελληνόφωνο Επιστημονικό Συνέδριο Εργαστηρίων Επικοινωνίας

Τόμ. 3, Αρ. 1 (2024)

3ο Ετήσιο Ελληνόφωνο Συνέδριο Εργαστηρίων Επικοινωνίας: Δημοσιογραφία, Μέσα και Επικοινωνία: Σύγχρονες προκλήσεις στην εποχή της Τεχνητής Νοημοσύνης



cclabs 2024

3ο Ετήσιο Ελληνόφωνο Συνέδριο Εργαστηρίων
Επικοινωνίας

Δημοσιογραφία, Μέσα και Επικοινωνία: Σύγχρονες
προκλήσεις στην εποχή της Τεχνητής Νοημοσύνης

29-30 Ιουνίου 2024, Θεσσαλονίκη

Διοργάνωση

Εργαστήρια: Εφαρμογών Πληροφορικής στα ΜΜΕ, Ηλεκτρονικών ΜΜΕ, του Τμήματος Δημοσιογραφίας & ΜΜΕ του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης



ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ



ΕΡΓΑΣΤΗΡΙΟ ΕΦΑΡΜΟΓΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΑ ΜΜΕ
ΤΜΗΜΑ ΔΗΜΟΣΙΟΓΡΑΦΙΑΣ
ΜΕΣΩΝ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΑΣ



Η Τεχνητή Νοημοσύνη στην Παραγωγή Πολυμέσων

Λάζαρος Βρύσης, Νικόλαος Βρύζας, Ελένη
Σιαμτανίδου, Χαράλαμπος Δημούλας

doi: [10.12681/cclabs.8068](https://doi.org/10.12681/cclabs.8068)

Copyright © 2025, Ετήσιο Ελληνόφωνο Επιστημονικό Συνέδριο
Εργαστηρίων Επικοινωνίας



Άδεια χρήσης [Creative Commons Αναφορά 4.0](https://creativecommons.org/licenses/by/4.0/).

Βιβλιογραφική αναφορά:

Βρύσης Λ., Βρύζας Ν., Σιαμτανίδου Ε., & Δημούλας Χ. (2025). Η Τεχνητή Νοημοσύνη στην Παραγωγή Πολυμέσων. *Ετήσιο Ελληνόφωνο Επιστημονικό Συνέδριο Εργαστηρίων Επικοινωνίας*, 3(1), 145-155.
<https://doi.org/10.12681/cclabs.8068>

Η Τεχνητή Νοημοσύνη στην Παραγωγή Πολυμέσων

Λάζαρος Βρύσης
ΕΔΙΠ, Τμήμα Δημοσιογραφίας & ΜΜΕ, ΑΠΘ
lvrysis@auth.gr

Νικόλαος Βρύζας
ΕΔΙΠ, Τμήμα Δημοσιογραφίας & ΜΜΕ, ΑΠΘ
nvryzas@auth.gr

Ελένη Σιαμαντίδου
Υποψήφια διδάκτωρ, Τμήμα Δημοσιογραφίας & ΜΜΕ, ΑΠΘ
siamtane@jour.auth.gr

Χαράλαμπος Δημούλας
Καθηγητής, Τμήμα Δημοσιογραφίας & ΜΜΕ, ΑΠΘ
babis@jour.auth.gr

Περίληψη

Η τεχνητή νοημοσύνη ενσωματώνεται ευρέως στην παραγωγή πολυμέσων, από την επεξεργασία εικόνας και ήχου έως τη δημιουργία βίντεο και παραγωγή μουσικής, ανατρέποντας παραδοσιακές μεθοδολογίες και διευρύνοντας τις δυνατότητες δημιουργίας περιεχομένου. Η έρευνα αυτή καταγράφει τα εργαλεία τεχνητής νοημοσύνης που χρησιμοποιούνται στην παραγωγή πολυμέσων και εξετάζει τις τεχνολογικές τους δυνατότητες, καθώς και την αποδοχή τους από το κοινό, με έμφαση στη συμβολή τους στη διαδραστικότητα και τη δημιουργικότητα. Η αυτοματοποίηση της επεξεργασίας βοηθά σημαντικά στη βελτίωση της αποτελεσματικότητας της οπτικοακουστικής παραγωγής, παρέχοντας εργαλεία που επιτρέπουν την εύκολη δημιουργία επαγγελματικού περιεχομένου, ανεξαρτήτως προηγούμενης εμπειρίας των χρηστών. Η ευρεία εφαρμογή της τεχνητής νοημοσύνης διευκολύνει τη δημιουργία και παρέχει πρόσβαση σε δημιουργικές διαδικασίες, ενώ η έρευνα και η αξιολόγηση των εργαλείων αυτών υποστηρίζουν την ανάγκη για πρότυπα που θα διασφαλίζουν δίκαιες και ηθικές πρακτικές στη χρήση της τεχνητής νοημοσύνης.

Λέξεις-κλειδιά: τεχνητή νοημοσύνη, επεξεργασία, μοντάζ, βίντεο, εικόνα, ήχος

1. Εισαγωγή

Η τεχνητή νοημοσύνη (TN) έχει φέρει επανάσταση στη δημιουργία και επεξεργασία οπτικοακουστικού περιεχομένου, προσφέροντας εργαλεία που ενισχύουν τη δημιουργικότητα και την αποτελεσματικότητα. Από αυτοματοποιημένη επεξεργασία εικόνων μέχρι προηγμένη σύνθεση ήχου και βίντεο, τα εργαλεία αυτά έχουν μετασχηματίσει τις παραδοσιακές μεθοδολογίες, ανοίγοντας νέους δρόμους για την παραγωγή περιεχομένου. Αυτή η εργασία έχει ως στόχο να προσφέρει μια σφαιρική εικόνα σχετικά με τα εργαλεία τεχνητής νοημοσύνης σε αυτόν τον τομέα, διερευνώντας τις τεχνολογικές δυνατότητες και τις ευκαιρίες που δημιουργούνται. Τα κριτήρια επιλογής των εργαλείων στηρίζονται στη συνάφεια τους με τις τρέχουσες εξελίξεις στην τεχνητή νοημοσύνη, την καινοτομία που εισάγουν στην παραγωγή πολυμέσων, καθώς και στην αποδοχή τους από το κοινό.

2. Μεθοδολογία

Η παρούσα έρευνα υιοθετεί μια θεωρητική προσέγγιση, η οποία βασίζεται στη σύνθεση της υπάρχουσας βιβλιογραφίας, για την εξερεύνηση των δυνατοτήτων και των εφαρμογών των εργαλείων τεχνητής νοημοσύνης σε διάφορους τομείς. Η μεθοδολογία επικεντρώνεται στη λεπτομερή απαρίθμηση και κατηγοριοποίηση των εργαλείων TN, εστιάζοντας στην παροχή παραδειγμάτων που αναδεικνύουν τις δυνητικές λειτουργικότητες αυτών των τεχνολογιών.

Τα δεδομένα για αυτή την έρευνα συλλέχθηκαν μέσω εκτεταμένης επισκόπησης της τρέχουσας βιβλιογραφίας και των υπάρχοντων εφαρμογών των εργαλείων τεχνητής νοημοσύνης σε διάφορους τομείς. Κάθε κατηγορία εργαλείων εξερευνήθηκε για να κατανοηθούν οι μοναδικές της συνεισφορές και περιορισμοί. Η αναλυτική προσέγγιση της έρευνας βασίζεται σε επαγωγική συλλογιστική, χωρίς τη χρήση εμπειρικών μεθόδων όπως ερωτηματολόγια. Επιπλέον, στηρίζεται στην αναθεώρηση της λογοτεχνίας και την προσωπική αξιολόγηση των εργαλείων TN για να κατανοήσει την απόδοση και την καταλληλότητά τους σε διάφορες εργασίες. Η ανάλυση περιλαμβάνει την κριτική των εργαλείων με βάση τη λειτουργικότητα, τη χρηστικότητα, την ολοκλήρωση στα υπάρχοντα συστήματα και τον αντίκτυπο στον αντίστοιχο τομέα. Αυτή η προσέγγιση επιτρέπει την παροχή μιας υποκειμενικής αλλά καλά ενημερωμένης άποψης σχετικά με την κατάσταση της TN και την εφαρμογή της στην παραγωγή πολυμέσων, συμπεριλαμβάνοντας τα συμπεράσματα από παραπλήσιες εργασίες και άρθρα.

3. Αποτελέσματα

3.1. Δημιουργία και επεξεργασία εικόνας

Στον χώρο της δημιουργίας και επεξεργασίας εικόνων, οι καινοτομίες που βασίζονται στην τεχνητή νοημοσύνη περιλαμβάνουν αυτοματοποιημένα εργαλεία βελτίωσης εικόνας, όπως η TN του Adobe Lightroom (Christensen, 2023), η οποία προσαρμόζει την ισορροπία χρωμάτων και αφαιρεί τον θόρυβο για να βελτιώσει την ποιότητα των φωτογραφιών και το Auto Tone του Adobe Photoshop (AI Photo Editing With Photoshop - Online & Desktop - Adobe, n.d.), το οποίο βελτιστοποιεί τη φωτεινότητα και την αντίθεση. Για τη μείωση και την όξυνση του θορύβου, το DeNoise AI της Topaz Labs ειδικεύεται στην αφαίρεση θορύβου διατηρώντας παράλληλα τις λεπτομέρειες (DeNoise AI - Remarkable Image Noise Reduction, 2024). Το DeNoise AI βελτιώνει την ποιότητα των εικόνων αφαιρώντας τον ανεπιθύμητο θόρυβο χωρίς να χάνει λεπτομέρειες, καθιστώντας τις φωτογραφίες πιο καθαρές και ευκρινείς. Είναι ιδιαίτερα χρήσιμο για φωτογραφίες που έχουν ληφθεί σε χαμηλό φωτισμό ή με υψηλό ISO.

Συνεχίζοντας, η αναγνώριση και η επεξεργασία αντικειμένων και χαρακτηριστικών του φυσικού κόσμου υποστηρίζεται από εργαλεία όπως το Face-Aware Liquify του Adobe Photoshop, το οποίο προσαρμόζει τα χαρακτηριστικά του προσώπου με ακρίβεια και το remove.bg, το οποίο χρησιμοποιεί AI για την ακριβή αφαίρεση φόντου (Adobe Creative Cloud | Adjust and exaggerate facial features, n.d.). Οι χρήστες μπορούν να ανεβάσουν μια εικόνα και το εργαλείο θα διαχωρίσει το κύριο θέμα από το φόντο, παρέχοντας μια εικόνα με διαφανές φόντο που μπορεί να χρησιμοποιηθεί για διάφορες εφαρμογές, όπως γραφικά, παρουσιάσεις και διαφημίσεις.

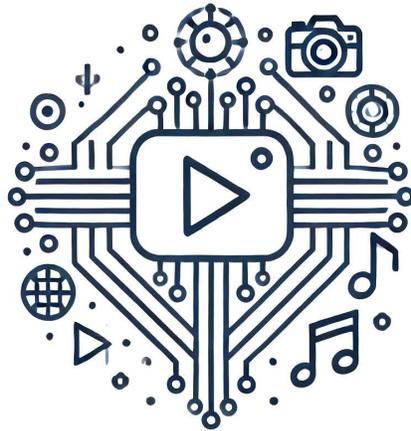
Τα φίλτρα και τα εφέ που υποστηρίζονται από AI περιλαμβάνουν τη μεταφορά στυλ (style transfer) του DeepArt, η οποία μετατρέπει τις φωτογραφίες σε έργα τέχνης στα στυλ διάσημων ζωγράφων και τα νευρωνικά φίλτρα του Photoshop που εφαρμόζουν πολύπλοκες επεξεργασίες υφής (Deep Art Effects Developer: Deep Art API for Enterprise Neural Style Transfer, n.d.). Για την αυτόματα διόρθωση διαφόρων περιοχών μιας εικόνας, το πινέλο διόρθωσης και το γέμισμα με επίγνωση περιεχομένου του Photoshop επιδιορθώνουν και γεμίζουν έξυπνα περιοχές χρησιμοποιώντας μοτίβα από το περιβάλλον περιεχόμενο.



Εικόνα 1: Στην πάνω εικόνα φαίνεται μια φωτογραφία με ορατό ψηφιακό θόρυβο. Στην κάτω αριστερή πλευρά παρουσιάζεται το αποτέλεσμα μετά την αποθουροποίηση μέσω τυπικών εφαρμογών λογισμικού (συγκεκριμένα από το Adobe Lightroom), ενώ στη δεξιά πλευρά μετά από αποθουροποίηση με το DeNoise AI. Είναι εύκολα αντιληπτό ότι στην κλασική αποθουροποίηση η εικόνα εμφανίζεται θολή, ενώ το DeNoise AI επιτυγχάνει την καταστολή του θορύβου διατηρώντας την οξύτητα της εικόνας.

Το Microsoft Designer αποτελεί ένα εργαλείο σχεδίασης που χρησιμοποιεί τεχνητή νοημοσύνη για να βοηθήσει τους χρήστες να δημιουργούν επαγγελματικά σχέδια (Narayanaswamy, 2024). Είναι ιδανικό για τη δημιουργία γραφικών για κοινωνικά δίκτυα, παρουσιάσεις και διαφημιστικά υλικά, προσφέροντας μια εύκολη και φιλική προς το χρήστη διεπαφή που διευκολύνει την πρόσβαση σε επαγγελματικής ποιότητας σχεδιασμό. Το DALL-E της OpenAI είναι ένα προηγμένο σύστημα τεχνητής νοημοσύνης που παράγει εικόνες

βασισμένες σε περιγραφές κειμένου (Ramesh et al., 2022). Αυτό το εργαλείο επιτρέπει στους χρήστες να μετατρέπουν τις λεκτικές τους ιδέες σε οπτικές απεικονίσεις, δημιουργώντας πολυμέσα που μπορούν να χρησιμοποιηθούν για εκπαιδευτικούς, καλλιτεχνικούς ή εμπορικούς σκοπούς. Το Midjourney είναι μια εφαρμογή τεχνητής νοημοσύνης που δημιουργεί εικόνες υψηλής ποιότητας από περιγραφές κειμένου. Παρόμοιο με το DALL-E, το Midjourney μετατρέπει τις περιγραφικές εντολές σε οπτικά έργα, επιτρέποντας την εξατομίκευση εικόνων για διάφορες χρήσεις όπως την ψηφιακή διαφήμιση, τον καλλιτεχνικό σχεδιασμό και άλλες οπτικές απεικονίσεις (Derevyanko, & Zalevska, 2023).



Εικόνα 2: Λογότυπο που δημιουργήθηκε από το DALL-E, χρησιμοποιώντας την προτροπή “Φτιάξε μου ένα λογότυπο για το έγγραφό μου με τίτλο ‘Η Τεχνητή Νοημοσύνη στην Παραγωγή Πολυμέσων’”.

3.2. Επεξεργασία ήχου και σύνθεση μουσικής

Στον ήχο και τη μουσική, οι εφαρμογές AI έχουν μεταμορφώσει τη σύνθεση και την επεξεργασία. Η Amper Music πρωτοπόρησε στη δημιουργία μουσικών κομματιών από το μηδέν και στην προσαρμογή τους σε διάφορα είδη και διαθέσεις. Υπηρεσία AI για τη σύνθεση, την παραγωγή και την ερμηνεία μουσικής. Αναπτύχθηκε από μια ομάδα επαγγελματιών μουσικών και ειδικών της τεχνολογίας και είναι το πρώτο AI που συνέθεσε και παρήγαγε ένα ολόκληρο μουσικό άλμπουμ (Taryn Southern, 2017). Τα ίχνη της υπηρεσίας αγνοούνται, αλλά, πλέον, είναι σίγουρο πως το concept τους ήταν εφικτό.

Και ήταν εφικτό γιατί πλέον υπάρχουν υπηρεσίες σαν την AIVA και το Suno.ai. Η AIVA ειδικεύεται στη δημιουργία μουσικών κομματιών που προσαρμόζονται σε διαφορετικά είδη (AIVA, the AI Music Generation Assistant, n.d.). Ένα από τα πιο εντυπωσιακά εργαλεία στον τομέα της τεχνητής νοημοσύνης είναι το Suno.ai, το οποίο ξεχωρίζει για τις εξαιρετικές του δυνατότητες στη δημιουργία και επεξεργασία ήχου. Το Suno.ai χρησιμοποιεί τεχνητή

νοημοσύνη για τη σύνθεση ήχου, τη δημιουργία μουσικής και φωνής με φυσικό ήχο. Η ευκολία χρήσης και η υψηλή ποιότητα των αποτελεσμάτων καθιστούν το εν λόγω εργαλείο ιδανικό για δημιουργούς περιεχομένου, μουσικούς και παραγωγούς που θέλουν να ενσωματώσουν ηχητικά στοιχεία υψηλής ποιότητας στις δημιουργίες τους (Chatterjee, 2024).

Επιπλέον, τα LANDR και το iZotope RX είναι κορυφαία εργαλεία επεξεργασίας ήχου, προσφέροντας αυτοματοποιημένη μίξη, mastering και αποκατάσταση ήχου χρησιμοποιώντας TN (LANDR, 2023). Το LANDR είναι μια πλατφόρμα προσφέρει εργαλεία για τη δημιουργία, την επεξεργασία και την κυκλοφορία μουσικής εκμεταλλευόμενη αλγορίθμους TN για να κάνει mastering σε κομμάτια μουσικής, εξασφαλίζοντας ότι ακούγονται επαγγελματικά. Το iZotope RX είναι το industry standard λογισμικό για την αποκατάσταση ήχου πλέον βασίζεται σε αλγορίθμους που κάνουν χρήση AI.

Οι τεχνολογίες σύνθεσης και τροποποίησης φωνής, όπως το Google Cloud Text-to-Speech, μετατρέπουν κείμενο σε ομιλία με φυσικό ήχο, ενώ το Descript επιτρέπει την κλωνοποίηση φωνής και τροποποιήσεις με απλή επεξεργασία κειμένου, επεκτείνοντας τις δημιουργικές δυνατότητες στην παραγωγή ήχου (All-in-One Video Editing, As Easy as a Doc, 2024). Ένα από τα χαρακτηριστικά γνωρίσματα του Descript είναι η ικανότητά του να μεταγράφει αυτόματα περιεχόμενο ήχου σε κείμενο και το αντίστροφο. Αυτή η μεταγραφή δεν είναι μόνο ακριβής, αλλά επιτρέπει την εύκολη επεξεργασία των αρχείων πολυμέσων με απλή επεξεργασία του κειμένου — η αποκοπή μιας λέξης στη μεταγραφή, για παράδειγμα, θα κόψει τον αντίστοιχο ήχο στη γραμμή χρόνου. Επιπλέον, το Descript προσφέρει μια δυνατότητα που ονομάζεται Overdub, η οποία επιτρέπει στους χρήστες να δημιουργούν μια συνθετική φωνή που ταιριάζει με τη φωνή ενός ατόμου. Με το Overdub, οι χρήστες μπορούν να πληκτρολογήσουν νέο περιεχόμενο ήχου που θα εκφωνήσει η επιλεγμένη φωνή, καθιστώντας το απίστευτα χρήσιμο για τη διόρθωση λαθών στον εγγεγραμμένο ήχο χωρίς να χρειάζεται επανεγγραφή. Επίσης, αξίζει να αναφερθεί το ElevenLabs Speech Synthesis, το οποίο εντυπωσιάζει με τη φυσικότητα της παραγόμενης ομιλίας, ακόμη και σε γλώσσες που δεν είναι τόσο δημοφιλείς, όπως η ελληνική (AI Voice Generator & Text to Speech, n.d.).

Πίνακας 1: Παραδείγματα παραγόμενου περιεχομένου από διάφορα εργαλεία αυτοματοποιημένης επεξεργασίας και παραγωγής ηχητικού περιεχομένου

| | Σύνδεσμος |
|-----------------------------|---|
| AIVA | https://research.playcompass.com/cclabs2024/aiva |
| Descript | https://research.playcompass.com/cclabs2024/descript |
| ElevenLabs Speech Synthesis | https://research.playcompass.com/cclabs2024/elevenlabs |
| iZotope RX | https://research.playcompass.com/cclabs2024/izotoperx |
| Suno AI | https://research.playcompass.com/cclabs2024/sunoai |

3.3. Παραγωγή βίντεο

Όσον αφορά τον τομέα της παραγωγής βίντεο, σημειώνεται πως και αυτός επωφελείται σημαντικά από την τεχνητή νοημοσύνη, με εργαλεία όπως το Magisto και το Auto Reframe του Adobe Premiere Pro να απλοποιούν την επεξεργασία και να διευκολύνουν τεχνικές, όπως το καδράρισμα. Το Magisto χρησιμοποιεί AI για ανίχνευση της σκηνής και την έξυπνη επεξεργασία, συναρμολογώντας κλιπ σε συνεκτικά βίντεο, ενώ το Auto Reframe προσαρμόζει αυτόματα τα βίντεο ώστε να ταιριάζουν σε διαφορετικές πλατφόρμες κοινωνικών μέσων (Automatically Reframe Video for Social Media Channels, n.d.). Στο post-production, το Autodesk Flame και το DaVinci Resolve φέρνουν οπτικά εφέ που βασίζονται σε AI και αυτοματοποιημένη χρωματική διαβάθμιση, αντίστοιχα (Blackmagic Design, n.d.). Η Bitmovin χρησιμοποιεί AI για βελτιστοποίηση κωδικοποίησης βίντεο ενώ το Watson Media της IBM παρέχει αναλυτικά στοιχεία περιεχομένου σε πραγματικό χρόνο για να ενισχύσει την αφοσίωση των θεατών (Smole, 2023).

Οι πλατφόρμες Synthesia και HeyGen αναπτύσσουν καινοτόμες λύσεις στον τομέα της δημιουργίας βίντεο με χρήση AI-generated avatars (Miumiu, 2024). Αυτές οι πλατφόρμες επιτρέπουν στους χρήστες να δημιουργούν παρουσιάσεις βίντεο με εικονικούς παρουσιαστές που μιλούν και αλληλεπιδρούν με το κοινό. Ο χρήστης μπορεί να επιλέξει ένα avatar, να εισάγει το κείμενο που επιθυμεί και η πλατφόρμα παράγει ένα βίντεο όπου το avatar απεικονίζει το κείμενο σε ομιλία. Τα εργαλεία αυτά είναι ιδιαίτερα χρήσιμα για τη δημιουργία εκπαιδευτικών βίντεο, εταιρικών παρουσιάσεων, διαφημίσεων και άλλων ψηφιακών περιεχομένων, καθώς προσφέρουν μια επαγγελματική εμφάνιση χωρίς την

ανάγκη για φυσικά στούντιο ή εκτεταμένο εξοπλισμό παραγωγής. Τέτοιες λύσεις αποτελούν ένα εξαιρετικό παράδειγμα του πώς η τεχνητή νοημοσύνη μπορεί να διευκολύνει και να επιταχύνει την παραγωγή ποιοτικού περιεχομένου, κάνοντας τη διαδικασία πιο αποδοτική και προσβάσιμη.

Τέλος, υπάρχουν επίσης πλατφόρμες όπως το Runway, το Dream Machine της Luma Labs, το Pika Art, που επιτρέπουν τη δημιουργία νέων κινούμενων εικόνων από το μηδέν με βάση απλές προτροπές (Marocsik, 2024). Αυτά τα εργαλεία ανοίγουν νέους δρόμους για τη δημιουργία περιεχομένου, παρέχοντας στους δημιουργούς τη δυνατότητα να εξερευνήσουν καινοτόμες ιδέες. Αυτά τα εργαλεία ανοίγουν νέους δρόμους για τη δημιουργία περιεχομένου, παρέχοντας στους δημιουργούς τη δυνατότητα να εξερευνήσουν καινοτόμες ιδέες. Ο χρήστης μπορεί να εισάγει μια προτροπή ή να τη συνοδεύσει με μια εικόνα της επιλογής του, ώστε να χρησιμοποιηθεί ως βάση για το τελικό αποτέλεσμα. Ωστόσο, όπως φαίνεται και από τα παραδείγματα που παρουσιάζονται στον Πίνακα 2, τα αποτελέσματα σε αυτήν την κατηγορία παραγωγής υλικού δεν είναι ακόμη ικανοποιητικά, καθώς η τεχνολογία βρίσκεται σε πρώιμο στάδιο και χρειάζεται βελτιώσεις για να ανταποκριθεί στις υψηλές απαιτήσεις της δημιουργικής βιομηχανίας.

Πίνακας 2: Παραδείγματα παραγόμενου περιεχομένου από διάφορα εργαλεία αυτοματοποιημένης δημιουργίας βίντεο

| | Σύνδεσμος |
|-------------------------|---|
| HeyGen | https://research.playcompass.com/cclabs2024/heygen |
| Luma Labs Dream Machine | https://research.playcompass.com/cclabs2024/dreammachine |
| Pika Art | https://research.playcompass.com/cclabs2024/pikaart |

3.4. Ολιστική παραγωγή πολυμέσων

Σε ευρύτερο πλαίσιο, εργαλεία σαν το Steve AI αξιοποιούν πολλούς από τους αυτοματισμούς που προαναφέρθηκαν για την end-to-end παραγωγή πολυμεσικού περιεχομένου (AI Video Generator - Create Videos From Text and Voice, n.d.). Για παράδειγμα, το εν λόγω εργαλείο δίνει τη δυνατότητα αυτόματης παραγωγής κινούμενων σχεδίων με μουσική, αφήγηση και υπότιτλους, παρέχοντας μόνο το σενάριο σε μορφή κειμένου. Προφανώς, το σενάριο μπορεί να παραχθεί με τη βοήθεια chatbots, όπως το ChatGPT ή το Gemini (Drapkin, 2024). Είναι ξεκάθαρο, πως έχουμε να κάνουμε με μια πλήρως αυτοματοποιημένη διαδικασία παραγωγής βίντεο, προσδιορίζοντας μόνο τον τίτλο του.

Πρέπει να επισημανθεί ότι τα εργαλεία αυτά εστιάζουν στη συναρμογή ή την επαναχρησιμοποίηση υπαρχόντων κλιπ για τη δημιουργία νέων βίντεο, αντί της δημιουργίας κινούμενης εικόνας από την αρχή.

Η πλατφόρμα Steve AI επιτρέπει στους χρήστες να μετατρέψουν ένα απλό κείμενο σε πλήρως αναπτυγμένο animation, χρησιμοποιώντας μια σειρά από προκαθορισμένους χαρακτήρες και σκηνές. Είναι ιδανικό για τη δημιουργία εκπαιδευτικών βίντεο, εταιρικών παρουσιάσεων και διαφημιστικών καμπανιών, προσφέροντας μια εύκολη και γρήγορη λύση για τη δημιουργία ποιοτικού οπτικού περιεχομένου χωρίς την ανάγκη για προηγμένες γνώσεις γραφιστικής ή βιντεοσκόπησης. Το VEED είναι μια διαδικτυακή πλατφόρμα επεξεργασίας βίντεο που προσφέρει εργαλεία για γρήγορη και αποτελεσματική δημιουργία και επεξεργασία βίντεο. Χρησιμοποιεί AI για την αυτόματη παραγωγή υποτίτλων, το καθάρισμα για διαφορετικές πλατφόρμες κοινωνικών μέσων και την ενίσχυση της ποιότητας του ήχου και του βίντεο. Επίσης, προσφέρει επιλογές συνεργασίας, επιτρέποντας σε πολλούς χρήστες να εργαστούν πάνω στο ίδιο έργο ταυτόχρονα. Το VEED είναι ιδιαίτερα δημοφιλές για την εύκολη χρήση του και την ικανότητά του να εξοικονομεί χρόνο στην παραγωγή περιεχομένου, καθιστώντας το ένα πολύτιμο εργαλείο για δημιουργούς περιεχομένου σε όλο τον κόσμο. Και τα δύο εργαλεία, VEED και Steve AI, είναι εκπρόσωποι μιας νέας γενιάς υπηρεσιών που χρησιμοποιούν την τεχνητή νοημοσύνη για να καταστήσουν τη δημιουργία ψηφιακού περιεχομένου πιο προσβάσιμη και αποδοτική.

Τέλος, ο ρόλος της τεχνητής νοημοσύνης στην προσβασιμότητα επισημαίνεται από υπηρεσίες όπως οι αυτόματοι υπότιτλοι του YouTube, το οποίο χρησιμοποιεί αναγνώριση φωνής για την παραγωγή υποτίτλων, ενισχύοντας την προσβασιμότητα (Use Automatic Captioning - YouTube Help, n.d.). Οι αυτοματισμοί που προσφέρει η τεχνητή νοημοσύνη για την επεξεργασία εικόνας, την παραγωγή ήχου και μουσικής και την επεξεργασία βίντεο όχι μόνο αυξάνουν την αποτελεσματικότητα και την ποιότητα της παραγωγής μέσω, αλλά εκδημοκρατίζουν επίσης την πρόσβαση σε δημιουργικές διαδικασίες υψηλού επιπέδου, επιτρέποντας στους δημιουργούς σε όλα τα επίπεδα να παράγουν αποτελεσματικά περιεχόμενο υψηλής ποιότητας.

Πίνακας 3: Παραδείγματα παραγόμενου περιεχομένου από διάφορα εργαλεία αυτοματοποιημένης οπτικοακουστικής παραγωγής

| | Σύνδεσμος |
|----------|---|
| Moovly | https://research.playcompass.com/cclabs2024/moovly |
| Steve AI | https://research.playcompass.com/cclabs2024/steveai |
| VEED | https://research.playcompass.com/cclabs2024/veedio |

4. Συμπεράσματα και προεκτάσεις

Η τεχνητή νοημοσύνη αποδεικνύεται ικανή να αυτοματοποιήσει την επεξεργασία πολυμέσων, συνεισφέροντας σημαντικά στη διαδικασία παραγωγής οπτικοακουστικού υλικού, προσθέτοντας αυτόματα υπότιτλους και ενσωματώνοντας γραφικά και εφέ, καθιστώντας την οπτικοακουστική παραγωγή πιο αποδοτική και λιγότερο χρονοβόρα. Επιπλέον, η τεχνητή νοημοσύνη βοηθά στην πρωτογενή δημιουργία οπτικοακουστικού υλικού, παρόλο που ακόμα η διαδικασία μπορεί να μην επιτυγχάνει πάντα το επιθυμητό αποτέλεσμα με την ακρίβεια που θα ήθελε ο χρήστης. Για παράδειγμα, η παραγωγή εικόνας ή λογοτύπου μπορεί να μην ανταποκρίνεται ακριβώς στις οδηγίες και οι φωνές από τις υπηρεσίες speech-to-text μπορεί να μην ακούγονται απόλυτα φυσικές.

Η εξάπλωση των εργαλείων τεχνητής νοημοσύνης εκδημοκρατίζει την παραγωγή μέσων, επιτρέποντας σε ιδιώτες και μικρά στούντιο να παράγουν περιεχόμενο με ποιότητα και ταχύτητα που παραδοσιακά ήταν δυνατή μόνο από μεγάλες εταιρείες παραγωγής. Αυτή η μετατόπιση όμως προκαλεί ερωτήματα για τον μειωμένο ρόλο των παραδοσιακών δεξιοτήτων και τις δεοντολογικές επιπτώσεις του περιεχομένου που παράγεται από την τεχνητή νοημοσύνη. Η ανάγκη για διαμόρφωση προτύπων που θα διέπουν τη χρήση της τεχνητής νοημοσύνης και θα εξασφαλίζουν δίκαιες πρακτικές είναι πιο επιτακτική από ποτέ.

Η διερεύνηση των ροών εργασίας στον κλάδο της δημοσιογραφίας και της επικοινωνίας που θα μπορούσαν να επωφεληθούν από αυτά τα εργαλεία TN είναι ενδιαφέρουσα. Παράλληλα, η πιλοτική δοκιμή επιλεγμένων εργαλείων σε πιο συγκεκριμένες ομάδες χρηστών και η αξιολόγησή τους μέσω ποσοτικής έρευνας θα βοηθήσει στην παραμετροποίηση και προσαρμογή των εργαλείων για να καλύπτουν συγκεκριμένες ανάγκες του χώρου της δημοσιογραφίας και της επικοινωνίας.

Αναφορές

Adobe Creative Cloud | Adjust and exaggerate facial features. (n.d.). <https://is.gd/VDfhGa>
AI Photo Editing With Photoshop – Online & Desktop – Adobe. (n.d.). <https://is.gd/9eBptf>
AI Video Generator - Create Videos from Text and Voice. (n.d.). Steve.AI. <https://www.steve.ai/>
AI Voice Generator & Text to Speech. (n.d.). ElevenLabs. <https://elevenlabs.io/app/speech-synthesis/>
AIVA, the AI Music Generation Assistant. (n.d.). <https://www.aiva.ai/>
Automatically Reframe Video for Social Media Channels. (n.d.). <https://is.gd/vVSMP8>
Blackmagic Design. (n.d.). DaVinci Resolve 19 – What’s new | BlackMagic Design. <https://is.gd/JtQrKZ>

- Chatterjee, S. (2024, July 11). Suno AI Review: Is It Worth the Hype? *Create Unique Royalty Free Music That Elevates Your Story*. <https://www.beatoven.ai/blog/suno-ai-review/>
- Christensen, R. (2023, April 18). *New Adobe Lightroom AI Innovations Empower Everyone to Edit Like a Pro | Adobe Blog*. <https://is.gd/0IXBv0>
- Deep Art Effects Developer: Deep Art API for Enterprise Neural Style Transfer*. (n.d.). <https://developer.deeparteffects.com/>
- Derevyanko, N., & Zalevska, O. (2023). Comparative Analysis of Neural Networks Midjourney, Stable Diffusion, and DALL-E and Ways of Their Implementation in the Educational Process of Students of Design Specialities. *Scientific Bulletin of Mukachevo State University. Series "Pedagogy and Psychology"*, 9(3), 36-44.
- Descript | All-in-One Video Editing, As Easy as a Doc*. (2024). <https://www.descript.com/>
- DeNoise AI – Remarkable Image Noise Reduction*. (2024, February 26). <https://www.topazlabs.com/denoise-ai>
- Drapkin, A. (2024, March 13). *Google Gemini vs ChatGPT 2024: AI Chatbot Head-to-Head Test*. Tech.co. <https://tech.co/news/google-bard-vs-chatgpt>
- LANDR. (2023, March 20). *LANDR: The Best Music Production Software for Creators*. <https://www.landr.com/>
- Marocsik, L. (2024, September 13). Home AI Tools SME. AI Tools for Small Businesses. <https://is.gd/6TUwIj>
- Miumiu, K. (2024, March 7). Unleashing Creativity: A Comparative Analysis of Synthesia and HeyGen in AI Video Generation. *Medium*. <https://is.gd/KS8tff>
- Narayanaswamy, A. (2024). Working with Microsoft Designer. In *Microsoft Copilot for Windows 11: Understanding the AI-Powered Features in Windows 11* (pp. 163-204). Berkeley, CA: Apress.
- Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical Text-Conditional Image Generation With Clip Latents. *arXiv preprint arXiv:2204.06125*, 1(2), 3.
- Smole, M. (2023, January 4). A.I. Encoding Uses Machine Learning to Speed Up Processing and Improve Quality. *Bitmovin*. <https://is.gd/KF2kIC>
- Taryn Southern. (2017). *Break Free - The First AI-Composed Pop Song | Lyrics by Taryn Southern* [Video]. YouTube. <https://youtu.be/XUs6CznN8pw>
- Use Automatic Captioning - YouTube Help*. (n.d.). <https://is.gd/11Q4bt>