

Συνέδρια της Ελληνικής Επιστημονικής Ένωσης Τεχνολογιών Πληροφορίας & Επικοινωνιών στην Εκπαίδευση

(2024)

8ο Πανελλήνιο Επιστημονικό Συνέδριο «Ένταξη και Χρήση των ΤΠΕ στην Εκπαιδευτική Διαδικασία»

The image shows the cover of a book or proceedings. At the top left is the logo of the University of Thessaly (ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ). At the top right is the logo of the Hellenic Association of Information and Communication Technologies in Education (ΕΠΙΤΠΕ). The main title is '8ο Πανελλήνιο Επιστημονικό Συνέδριο Ένταξη και Χρήση των ΤΠΕ στην Εκπαιδευτική Διαδικασία' (8th Panhellenic Scientific Conference on the Integration and Use of ICT in the Educational Process). The location and dates are 'Βόλος, 27-29 Σεπτεμβρίου 2024'. Below the title, it lists the organizing institutions: Παιδαγωγικό Τμήμα Ειδικής Αγωγής, Παιδαγωγικό Τμήμα Προσχολικής Εκπαίδευσης, Παιδαγωγικό Τμήμα Δημοτικής Εκπαίδευσης, and Τμήμα Επιστήμης Φυσικής Αγωγής & Αθλητισμού. The editors are Χαράλαμπος Καραγιαννίδης, Ηλίας Καρασαββίδης, Βασίλης Κόλλας, and Μαρίνα Παπαστεργίου. The website is etpe2024.uth.gr and the ISBN is 978-618-5866-00-6.

Οικειοποίηση των διαδικασιών καθοδήγησης μοντέλων παραγωγικής τεχνητής νοημοσύνης, για τη δημιουργία συνθετικών εικόνων από παιδιά 12-14 ετών: Μια μελέτη περίπτωσης

Σοφία Τζήμα, Ηλίας Καρασαββίδης

Βιβλιογραφική αναφορά:

Τζήμα Σ., & Καρασαββίδης Η. (2025). Οικειοποίηση των διαδικασιών καθοδήγησης μοντέλων παραγωγικής τεχνητής νοημοσύνης, για τη δημιουργία συνθετικών εικόνων από παιδιά 12-14 ετών: Μια μελέτη περίπτωσης. *Συνέδρια της Ελληνικής Επιστημονικής Ένωσης Τεχνολογιών Πληροφορίας & Επικοινωνιών στην Εκπαίδευση*, 275-289. ανακτήθηκε από <https://eproceedings.epublishing.ekt.gr/index.php/cetpe/article/view/8448>



Οικειοποίηση των διαδικασιών καθοδήγησης μοντέλων παραγωγικής τεχνητής νοημοσύνης, για τη δημιουργία συνθετικών εικόνων από παιδιά 12-14 ετών: Μια μελέτη περίπτωσης

Σοφία Τζήμα, Ηλίας Καρασαββίδης
sotzima@uth.gr, ikaras@uth.gr

Παιδαγωγικό Τμήματα Προσχολικής Εκπαίδευσης, Πανεπιστήμιο Θεσσαλίας

Περίληψη

Σκοπός αυτής της μελέτης ήταν να διερευνήσει πώς τρεις ομάδες παιδιών ηλικίας 12-14 ετών δημιούργησαν συνεργατικά προτροπές, για την κατασκευή συνθετικών εικόνων με τη χρήση της παραγωγικής τεχνητής νοημοσύνης. Συγκεκριμένα, η εργασία επικεντρώθηκε στις διαδικασίες και στρατηγικές που ανέπτυξαν και ακολούθησαν τα παιδιά για τη σύνταξη προτροπών, αρχικά αυτόνομα, στη συνέχεια με σταδιακή καθοδήγηση και τέλος με μετα-έλεγχο των αρχών που υιοθέτησαν. Επιπλέον, η μελέτη εξέτασε το είδος των προτροπών που τα παιδιά δημιούργησαν, καθώς και τον τρόπο που διαχειρίστηκαν και ενσωμάτωσαν τις παρεχόμενες αρχές προτροπών. Το ερευνητικό σχέδιο που υιοθετήθηκε ήταν η περιγραφική μελέτη περίπτωσης με ποιοτική προσέγγιση. Τα αποτελέσματα της μελέτης έδειξαν ότι αρκετές αρχές προτροπών αναπτύχθηκαν διαισθητικά από τα παιδιά, χωρίς ρητή καθοδήγηση. Στο πλαίσιο των αλληλεπιδράσεων με το μοντέλο DALL-E, τα παιδιά αντιμετώπισαν δυσκολίες με τις προτροπές και τις παραγόμενες εικόνες. Ωστόσο, στην προσπάθεια υπέρβασης των δυσκολιών, τα παιδιά πειραματίστηκαν αναπτύσσοντας διάφορες στρατηγικές.

Λέξεις κλειδιά: Παραγωγική τεχνητή νοημοσύνη, Συνεργασία παιδιών-GenAI, Συνθετικές εικόνες, Prompt engineering

Εισαγωγή

Τα τελευταία πενήντα χρόνια, όταν μια υποτιθέμενη μοναδική ανθρώπινη ικανότητα ή δεξιότητα αυτοματοποιείται μέσω της τεχνολογίας την αποκαλούμε τεχνητή νοημοσύνη «AI». Ωστόσο, μόλις αυτός ο αυτοματισμός είναι απρόσκοπτα και πλήρως επιτυχημένος, τείνουμε να σταματάμε να τον περιγράφουμε ως τεχνητή νοημοσύνη. Η ιστορία των συστημάτων και εργαλείων ψηφιακών μέσων είναι γεμάτη από τέτοιες «στιγμές τεχνητής νοημοσύνης» – εκπληκτικές στην αρχή, στη συνέχεια όμως θεωρούνται δεδομένες και ξεχασμένες ως «AI» (Manovich et al., 2023).

Σύμφωνα με τους Manovich et al. (2023) και Orpenlaender (2023), η νέα επανάσταση της γενετικής τεχνητής νοημοσύνης φαίνεται να είναι εξίσου σημαντική, όσο η εφεύρεση της φωτογραφίας τον δέκατο ένατο αιώνα. Οι όροι «γενετικά μέσα», «παραγωγικά μέσα τεχνητής νοημοσύνης», «γενετική τεχνητή νοημοσύνη», «συνθετικά μέσα» και «GAI» είναι όλοι ισοδύναμοι και αναφέρονται στη διαδικασία δημιουργίας περιεχομένου νέων μέσων με βαθιά νευρωνικά δίκτυα, όπως εικόνες, κινούμενα σχέδια, βίντεο, κείμενο, μουσική, τρισδιάστατα μοντέλα κ.α. Αυτά τα δίκτυα εκπαιδεύονται σε τεράστιες συλλογές αντικειμένων και μέσων, που υπάρχουν ήδη στον παγκόσμιο ιστό. Οι δημοφιλείς τύποι τεχνητών νευρωνικών δικτύων, για τη δημιουργία πολυμέσων, περιλαμβάνουν διάφορα μοντέλα (π.χ. διάχυσης, κειμένου σε εικόνα, δίκτυα ανταγωνιστικής παραγωγής και μετασχηματιστές). Στη δημιουργία σταθερών και κινούμενων εικόνων με χρήση νευρωνικών

δικτύων, οι όροι δημιουργία εικόνων, συνθετική εικόνα, εικόνα τεχνητής νοημοσύνης και οπτικές τεχνολογίες τεχνητής νοημοσύνης μπορούν να χρησιμοποιηθούν εναλλακτικά (Manovich et al., 2023).

Με τα παραγωγικά μέσα τεχνητής νοημοσύνης, εμφανίζεται μια ριζικά νέα μέθοδος. Οι υπολογιστές χρησιμοποιούν μεγάλα σύνολα δεδομένων, από υπάρχουσες αναπαραστάσεις σε διάφορα μέσα, για να προβλέψουν νέες εικόνες (σταθερές και κινούμενες). «Πρόβλεψη» είναι ο πραγματικός όρος που χρησιμοποιείται συχνά από ερευνητές γενετικής τεχνητής νοημοσύνης, για την περιγραφή μεθόδων οπτικής παραγωγής μέσω (Manovich et al., 2023). Όταν εργαζόμαστε με ένα μοντέλο γενετικής τεχνητής νοημοσύνης κειμένου σε εικόνα, το νευρωνικό δίκτυο προσπαθεί να προβλέψει τις εικόνες που αντιστοιχούν καλύτερα στο κείμενο που εισάγεται. Το αποτέλεσμα αυτής της διαδικασίας είναι ότι το μοντέλο μπορεί να εξαγάγει εικόνες με το να μεταφράσει-ταιριάζει οπτικά τις εισαγωγές κειμένου που δίνονται από τους χρήστες. Ορισμένες από αυτές τις μεταφράσεις μπορούν πλέον να εκτελούνται αυτόματα σε μαζική κλίμακα, χάρη στα τεχνητά νευρωνικά δίκτυα, αποτελώντας ένα νέο μέσο επικοινωνίας και δημιουργίας πολιτισμού. Αυτό που κάποτε ήταν μια εξειδικευμένη καλλιτεχνική πράξη τώρα είναι μια τεχνολογική ικανότητα διαθέσιμη σε όλους. Με άλλα λόγια, αναπτύχθηκε μια τεχνολογία που, από άποψη πολυπλοκότητας, μοιάζει εξαιρετικά με τον ανθρώπινο εγκέφαλο. Δεν κατανοούμε πλήρως πώς λειτουργεί η τεχνολογία GenAI, όπως δεν κατανοούμε πλήρως την ανθρώπινη διάνοια και δημιουργικότητα (Manovich et al., 2023).

Η τρέχουσα γενιά γενετικών συστημάτων κειμένου σε εικόνα (GAI text-to-image), μετά από εκπαίδευση σε πολύ μεγάλα και διαφορετικά σύνολα δεδομένων, που αποτελούνται από δισεκατομμύρια ή και τρισεκατομμύρια ζεύγη εικόνων και κειμένου, μπορούν να δημιουργήσουν συνθετικές εικόνες και γραφικά επαγγελματικού επιπέδου (Manovich et al., 2023). Πρόκειται για μια σημαντική και καινοτόμο τεχνολογία που επιτρέπει τη δημιουργία ρεαλιστικών συνθετικών εικόνων με βάση τις περιγραφές που παρέχονται από τους χρήστες, με τη χρήση φυσικής γλώσσας (Dodić & Čungurski, 2023). Η εφαρμογή της συναντάται σε τομείς όπως της βιομηχανίας ταινιών και βιντεοπαιχνιδιών, της τέχνης και του σχεδίου, της ιατρικής, της αρχιτεκτονικής και πολλών άλλων (Dodić & Čungurski, 2023).

Μερικά από τα πιο γνωστά γεννητικά μέσα (GAI) που έχουν σχεδιαστεί ειδικά για τη δημιουργία εικόνων από περιγραφές φυσικής γλώσσας (προτροπές), όπως το DALL-E, το Midjourney και το Stable Diffusion κ.α. είναι εύχρηστα και έχουν καταστήσει τη δημιουργία συνθετικών εικόνων προσβάσιμη σε όλους τους χρήστες αρχάριους και μη (Oppenlaender 2022). Αυτά τα παραγωγικά μέσα τεχνητής νοημοσύνης χρησιμοποιούν κείμενο ως είσοδο για τη σύνθεση εικόνας. Αυτή η είσοδος, γνωστή ως «προτροπή» ή «υπόδειξη» (prompt), περιγράφει την εικόνα στο σύστημα, το οποίο στη συνέχεια δημιουργεί αυτόνομα μία ή περισσότερες εικόνες. Η πρακτική της δημιουργίας προτροπών εισαγωγής αναφέρεται ως άμεση μηχανική ή προτροπή για συντομία (Oppenlaender et al., 2023). Σύμφωνα με τους Gu et al. (2023), η προτροπή μπορεί να οριστεί ως οι πρόσθετες πληροφορίες ή υποδείξεις που παρέχονται σε ένα μοντέλο για να καθοδηγήσουν τη συμπεριφορά του ή να το βοηθήσουν να εκτελέσει μια συγκεκριμένη εργασία. Η λειτουργία και τα αποτελέσματα της γενετικής τεχνητής νοημοσύνης δεν είναι προκαθορισμένα ή σταθερά, αλλά προκύπτουν μέσω αλληλεπιδράσεων με ανθρώπους, που καθοδηγούν το εργαλείο τεχνητής νοημοσύνης χρησιμοποιώντας φυσική γλώσσα (Vartiainen & Tedre, 2024). Αυτό σημαίνει πώς η γενετική τεχνητή νοημοσύνη δε διαθέτει μαγική ικανότητα να δημιουργεί συνθετικές εικόνες από μόνη της, αλλά πρέπει να χρησιμοποιείται από τον ανθρώπινο παράγοντα. Επίσης, συνεπάγεται ότι η φυσική γλώσσα έχει εξελιχθεί ώστε να έχει ισχυρό διαμεσολαβητικό αντίκτυπο στη δημιουργία εικόνων, χρησιμεύοντας ως εργαλείο για να καθοδηγήσει ένα γενετικό μέσο στην

παραγωγή συνθετικών εικόνων. Αυτή η διαμεσολαβούμενη δράση παραγωγής προτροπών διαμορφώνεται επίσης από το γενετικό μέσο, καθώς η προτροπή πρέπει να δημιουργηθεί για να επικοινωνεί αποτελεσματικά με το μοντέλο GAI, καθοδηγώντας το να παράγει εικόνες σε συγκεκριμένο στυλ, δεδομένη σύνθεση σκηνής, χρησιμοποιώντας ορισμένα υλικά, γωνία θέασης κ.α. (Vartiainen & Tedre, 2024). Μια πτυχή της προτροπής που σχετίζεται με τον μηχανικό της χαρακτήρα είναι ότι συχνά περιλαμβάνει συστηματικό πειραματισμό μέσω δοκιμής και πλάνης, έτσι ώστε οι χρήστες να μπορούν να επαναλάβουν και να βελτιώσουν τις ιδέες και τις δημιουργίες τους, τροποποιώντας τις προτροπές τους ή κάνοντας παραλλαγές για να έχουν το επιθυμητό αποτέλεσμα (Vartiainen & Tedre, 2024). Η πρόκληση για τον χρήστη δεν είναι μόνο να βρει τις σωστές λέξεις για να περιγράψει την επιδιωκόμενη έξοδο-εικόνα, αλλά και να προβλέψει πώς άλλα άτομα θα περιέγραφαν και θα αντιδρούσαν σε αυτό το αποτέλεσμα στον παγκόσμιο ιστό (Oppenlaender et al., 2023). Ωστόσο, οι χρήστες εξακολουθούν να αντιμετωπίζουν προκλήσεις όταν προσπαθούν να δημιουργήσουν εικόνες που αντικατοπτρίζουν συγκεκριμένα θέματα και τις μορφές τους, καθώς τα μοντέλα παραγωγής εικόνων, απαιτούν τις κατάλληλες προτροπές για να επιτύχουν τις επιθυμητές εικόνες (Kim et al., 2023).

Η έρευνα για την αλληλεπίδραση των ανθρώπων με τα παραγωγικά μοντέλα τεχνητής νοημοσύνης, όσον αφορά τη δημιουργία συνθετικών εικόνων, έχει επικεντρωθεί σε συγκεκριμένες ομάδες του γενικού πληθυσμού όπως καλλιτέχνες, φοιτητές και επαγγελματίες διαφόρων ειδικοτήτων (Fiebrink, 2019· Audry, 2021· Caramiaux & Fdili Alaoui, 2022). Ωστόσο, ελάχιστες είναι οι μελέτες που εντοπίζονται στην πληθυσμιακή ομάδα παιδιών 8-15 ετών και ακόμα λιγότερες στην προσχολική και πρώτη σχολική ηλικία (Vartiainen et al., 2023).

Σκοπός έρευνας

Σκοπός της παρούσας μελέτης ήταν η διερεύνηση των τρόπων με τους οποίους μαθητές/τριες (12-14 ετών) δημιούργησαν συνεργατικά προτροπές κειμένου, για την κατασκευή ψηφιακών συνθετικών εικόνων με τη χρήση της παραγωγικής τεχνητής νοημοσύνης (GAI). Συγκεκριμένα, η παρούσα μελέτη επικεντρώθηκε στις διαδικασίες και στρατηγικές που ανέπτυξαν και ακολούθησαν τα παιδιά για τη σύνταξη προτροπών, ώστε να παραγάγουν συνθετικές εικόνες αντίστοιχες με τις δοθείσες εικόνες αναφοράς. Αρχικά αυτόνομα, στη συνέχεια με σταδιακή καθοδήγηση και τέλος με μετα-έλεγχο των αρχών που υιοθέτησαν. Επίσης, η παρούσα μελέτη εστίασε στο είδος των κειμενικών προτροπών που δημιούργησαν τα παιδιά χωρίς καθοδήγηση, καθώς και του τρόπου που διαχειρίστηκαν και ενσωμάτωσαν τις παρεχόμενες αρχές σύνταξης προτροπών, για την κατασκευή αντίστοιχων εικόνων. Επομένως, βασική πρόκληση για τα παιδιά ήταν η επιλογή κατάλληλων κειμενικών προτροπών, μέσα από μια διαδικασία παρατήρησης, επεξεργασίας και ανάλυσης των συνθετικών εικόνων αναφοράς, καθώς και ενσωμάτωσης των αρχών σύνταξης, με στόχο τον σχεδιασμό αντίστοιχων εικόνων (Baidoo-Anu & Ansa, 2023). Επιπλέον, η μελέτη εξέτασε το πώς τα παιδιά ερμήνευσαν και αξιολόγησαν τις ψηφιακές συνθετικές εικόνες που κατασκεύασαν σε σχέση με τις εικόνες αναφοράς, καθώς και τις δυσκολίες που συνάντησαν. Τέλος, η μελέτη στόχευσε στην αποτίμηση της εμπειρίας των παιδιών και των αντιλήψεων για την αλληλεπίδρασή τους με το μοντέλο παραγωγικής τεχνητής νοημοσύνης DALL-E. Τα ερευνητικά ερωτήματα που καθοδήγησαν τη μελέτη ήταν τα εξής:

- 1) Με ποιους τρόπους τα παιδιά αλληλεπίδρασαν με τη διεπαφή του μοντέλου;
- 2) Ποιες στρατηγικές και αρχές κειμενικών προτροπών εφάρμοσαν αυτόνομα τα παιδιά, στο πλαίσιο αναπαραγωγής συνθετικών εικόνων αναφοράς;

- 3) Ποιες από τις παρεχόμενες αρχές σύνταξης προτροπών ενσωμάτωσαν τα παιδιά στις υφιστάμενες δικές τους;
- 4) Πώς τα παιδιά ερμήνευσαν και αξιολόγησαν εμπειρικά την αλληλεπίδρασή τους με το μοντέλο DALL-E, για τη δημιουργία συνθετικών εικόνων;

Μεθοδολογία

Το ερευνητικό σχέδιο που υιοθετήθηκε στην έρευνα ήταν αυτό της περιγραφικής μελέτης περίπτωσης με ποιοτική προσέγγιση. Πραγματοποιήθηκε δειγματοληψία ευκολίας, από την οποία συγκροτήθηκαν τρεις ομάδες. Η μία ομάδα αποτελούνταν από δύο παιδιά διαφορετικού φύλου, ενώ οι άλλες δύο ομάδες από δύο παιδιά του ίδιου φύλου. Οι έξι συμμετέχοντες/ουσες ήταν παιδιά ηλικίας 12-14 ετών και είχαν γνώσεις αγγλικών επιπέδου B2. Τα παιδιά, καθώς και οι γονείς τους ενημερώθηκαν και συναίνεσαν να λάβουν μέρος στη σχετική μελέτη. Από τη διερεύνηση των προηγούμενων αντιλήψεων και γνώσεων προέκυψε ότι κανένα παιδί δεν είχε προηγούμενη εμπειρία στη δημιουργία εικόνων με τη χρήση συνθετικών μέσων. Αντίθετα, σχεδόν όλα τα παιδιά ανέφεραν ότι είχαν ακούσει για το ChatGPT και ένα παιδί δήλωσε ότι το είχε χρησιμοποιήσει μία φορά για αναζήτηση πληροφοριών. Η διδακτική παρέμβαση διενεργήθηκε σε διαφορετική χρονική στιγμή για κάθε ομάδα, με τις ίδιες ακριβώς προδιαγραφές.

Διδακτική παρέμβαση

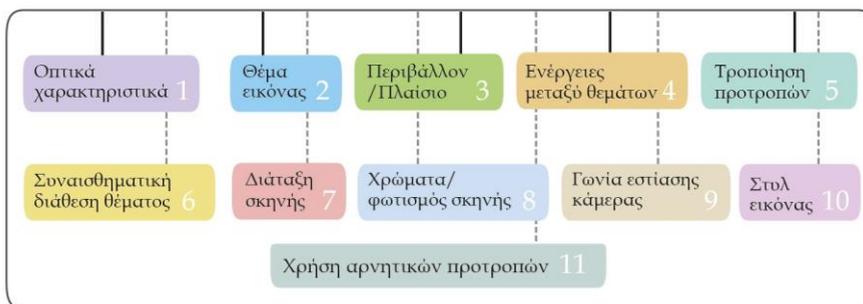
Το γενικό σχήμα της διδακτικής παρέμβασης παρουσιάζεται στο σχήμα 1. Αρχικά, η πρώτη συγγραφέας έκανε μία μικρή εισαγωγή σχετικά με τα παραγωγικά μέσα τεχνητής νοημοσύνης και το περιεχόμενο που παραγάγουν, εστιάζοντας σε αυτά που δημιουργούν συνθετικές εικόνες μέσω προτροπών σε φυσική γλώσσα. Παρουσιάστηκε στα παιδιά η διεπαφή του εργαλείου γενετικής τεχνητής νοημοσύνης DALL-E, καθώς και η διαδικασία χρήσης του, για τη δημιουργία εικόνων μέσω προτροπών. Συγκεκριμένα, παρουσιάστηκαν δύο συνθετικές εικόνες, καθώς και οι αντίστοιχες προτροπές τους. Μετά, προβλήθηκαν διαδοχικά (σε μία δεύτερη οθόνη Η/Υ) τρεις συνθετικές εικόνες-αναφοράς, χωρίς τις αντίστοιχες προτροπές. Τα παιδιά κλήθηκαν να αλληλεπιδράσουν με το μοντέλο και να συντάξουν τις πρώτες τους κειμενικές προτροπές με σκοπό να παραγάγουν τις αντίστοιχες εικόνες, χωρίς καμία καθοδήγηση. Η αλληλεπίδραση με το μέσο DALL-E μπορούσε να γίνει είτε πληκτρολογώντας, είτε εκφωνώντας τις αντίστοιχες προτροπές. Για κάθε προτροπή, το μοντέλο δημιουργούσε τέσσερις εικόνες και η κάθε ομάδα διάλεγε και αποθήκευε αυτή που θεωρούσε πως ταίριαζε περισσότερο με την εικόνα αναφοράς (Microsoft, 2024).



Σχήμα 1. Διδακτική παρέμβαση

Στη συνέχεια, με την προοδευτική παροχή αρχών καθοδήγησης προτροπών, τα παιδιά κλήθηκαν να γράψουν προτροπές, για να αναπαραγάγουν τέσσερις νέες συνθετικές εικόνες αναφοράς. Η παροχή αρχών καθοδήγησης προτροπών (σχήμα 2) προέκυψε από τη βιβλιογραφία και περιελάμβανε τα εξής: 1) Αναγνώριση των οπτικών χαρακτηριστικών της εικόνας αναφοράς και προσεκτική επιλογή των λέξεων-κλειδιών για την περιγραφή της

(Atlas, 2023 · Choi et al., 2023 · Ali et al., 2024). 2) Εντοπισμός του θέματος της εικόνας, με ειδική αναφορά σε αντικείμενα και χαρακτήρες (Taheri et al., 2023), καθώς και περιγραφή των στοιχείων τους (Venkataramanan, 2023). 3) Προσδιορισμός του περιβάλλοντος στο οποίο βρίσκεται το θέμα (Taheri et al., 2023 · Venkataramanan, 2023). 4) Περιγραφή των ενεργειών που εντοπίζονται στην εικόνα μεταξύ χαρακτήρων ή και αντικειμένων (Choi et al., 2023 · Taheri et al., 2023). 5) Επανασχεδιασμός-τροποποίηση των κειμενικών προτροπών με σκοπό τη συγχώνευση λέξεων-κλειδιών, για τη δημιουργία περιγραφικών αλλά και συνάμα συνοπτικών προτροπών (Atlas, 2023 · Choi et al., 2023 · Venkataramanan, 2023). 6) Περιγραφή της συναισθηματικής διάθεσης του θέματος και του περιβάλλοντος (Choi et al., 2023 · Venkataramanan, 2023). 7) Περιγραφή των σχέσεων διάταξης σκηνής της εικόνας μεταξύ υποκειμένων, αντικειμένων και ενεργειών (Choi et al., 2023 · Taheri et al., 2023). 8) Περιγραφή των χρωμάτων και του φωτισμού της σκηνής (Sanchez, 2023 · Venkataramanan, 2023). 9) Αναφορά της γωνίας εστίασης της κάμερας σε σχέση με την εικόνα αναφοράς (Sanchez, 2023). 10) Περιγραφή του στυλ της εικόνας όπως ελαιογραφία, κόμικς, 3D κ.α. (Sanchez, 2023 · Taheri et al., 2023 · Venkataramanan, 2023). 11) Χρήση αρνητικών προτροπών, που περιγράφουν οτιδήποτε δεν επιθυμούμε να περιλαμβάνει η συνθετική εικόνα (Atlas, 2023). Οι προτροπές δόθηκαν στα παιδιά σταδιακά ανά εικόνα αναφοράς με την παραπάνω σειρά και συγκεκριμένα: Οι προτροπές 1 και 2 δόθηκαν με την παρουσίαση της 4ης εικόνας αναφοράς, οι προτροπές 3, 4, 5 δόθηκαν με την παρουσίαση της 5ης εικόνας αναφοράς, οι προτροπές 6, 7, 8 δόθηκαν με την παρουσίαση της 6ης εικόνας αναφοράς, και τέλος οι προτροπές 9, 10, 11 δόθηκαν με την παρουσίαση της 7ης εικόνας αναφοράς. Τα παιδιά μετά από κάθε παροχή προτροπών ανά εικόνα αναφοράς ήταν ελεύθερα να ενεργήσουν κατά βούληση.



Σχήμα 2. Αρχές σύνταξης προτροπών

Στο τέλος, μετά την καθοδήγηση, ζητήθηκε από κάθε ομάδα να αναπαραγάγει αυτόνομα δύο εικόνες αναφοράς, με σκοπό την εξέταση της υιοθέτησης και ενσωμάτωσης αρχών για τη δημιουργία προτροπών. Μετά την παρέμβαση, ακολούθησε συνέντευξη κάθε ομάδας με τη χρήση ημιδομημένων ανοιχτών ερωτήσεων, για την καταγραφή της συνολικής εμπειρίας τους. Οι συνθετικές εικόνες αναφοράς που παρουσιάστηκαν στα παιδιά ήταν ανεξάρτητες μεταξύ τους, δηλαδή μπορούσαν να προσεγγιστούν αυτόνομα η καθεμία. Επίσης, παρουσιάστηκαν με μία συγκεκριμένη σειρά και είχαν βαθμό αύξουσας δυσκολίας όσον αφορά την περιπλοκότητα του θέματος. Ο συνολικός χρόνος της παρέμβασης ήταν 120 λεπτά και κατανεμήθηκε επιμέρους ως εξής: Εισαγωγή (10 λεπτά), αυτόνομη εκτέλεση-καθοδήγηση-μεταέλεγχος (90 λεπτά), διάλλειμα (5 λεπτά) και ομαδική συνέντευξη με κάθε ομάδα (15 λεπτά).

Συλλογή δεδομένων

Η έρευνα διεξήχθη τον Ιούνιο του 2024 και η συλλογή των δεδομένων έγινε με τη χρήση τριών κύριων μέσων: της καταγραφής οθόνης, ηχογράφησης και συνέντευξης. Ειδικότερα, πραγματοποιήθηκε καταγραφή οθόνης και παράλληλη ηχογράφηση, κατά τη διάρκεια της παρέμβασης, ώστε να αποτυπωθούν λεπτομερώς οι ενέργειες και οι συμπεριφορές των παιδιών κατά την αλληλεπίδραση τους με το παραγωγικό μέσο τεχνητής νοημοσύνης DALL-E. Σε δεύτερο χρόνο, έγινε αυτολεξεί απομαγνητοφώνηση και καταγραφή των σημαντικότερων σημείων. Μετά, το τέλος της παρέμβασης κάθε ομάδας, ακολούθησε ημιδομημένη συνέντευξη που ηχογραφήθηκε και στη συνέχεια απομαγνητοφωνήθηκε λέξη προς λέξη. Στο πλαίσιο αυτό σχεδιάστηκε ένας οδηγός συνέντευξης ενδεικτικά παρατίθενται κάποιες ερωτήσεις: «*Πώς σας φάνηκε η δημιουργία εικόνων με τη χρήση της γενετικής τεχνητής νοημοσύνης; Μπορείτε να μου πείτε πως σκεφτήκατε για να κάνετε τις προτροπές;*» κ.α.

Ανάλυση δεδομένων

Η ανάλυση δεδομένων βασίστηκε στην θεματική ανάλυση περιεχομένου με μια επαγωγική προσέγγιση. Οι κατηγορίες κωδικοποίησης αναπτύχθηκαν με βάση τα δεδομένα (Elo & Kyngäs, 2008). Τα μηνύματα προτροπών κάθε ομάδας αποθηκεύτηκαν στο συνθετικό μέσο DALL-E και εξετάστηκαν ως προς την εξελκτική πορεία ανάπτυξης διαδικασιών και στρατηγικών προτροπών, για την παραγωγή των συνθετικών εικόνων. Με βάση την ανασκόπηση των δεδομένων εντοπίστηκαν μοτίβα, που αντιπροσώπευαν βασικές αλλαγές στις σχέσεις που διαμορφώθηκαν, μέσω των αλληλεπιδράσεων με το εργαλείο παραγωγικής τεχνητής νοημοσύνης. Συγκεκριμένα τα θέματα που αναδείχθηκαν ήταν τα εξής:

- 1) Ανάπτυξη αυτόνομων διαδικασιών και στρατηγικών κειμενικών προτροπών από τα παιδιά, για την αναπαραγωγή συνθετικών εικόνων αναφοράς.
- 2) Διαχείριση συμβουλών που αφορούσαν τις αρχές καθοδήγησης, για την αναθεώρηση ή βελτίωση των υφιστάμενων προτροπών.
- 3) Διερεύνηση της κατανόησης, εφαρμογής και ενσωμάτωσης των παρεχόμενων αρχών μετά την καθοδήγηση. Συγκριτική ανάλυση προτροπών για τον εντοπισμό της συνολικής προόδου.
- 4) Αξιολόγηση της διαδικασίας αλληλεπίδρασης με το συνθετικό μέσο DALL-E.

Η διαδικασία ανάλυσης των δεδομένων απαιτούσε τη συνεχή μετακίνηση μεταξύ των παραπάνω τεσσάρων θεματικών, για την αναγνώριση των μεταβαλλόμενων σχεσιακών προτύπων και την ανάλυση της πορείας προτροπών που ακολούθησαν τα παιδιά.

Αποτελέσματα

Θέμα: Διεπαφή-τρόπος αλληλεπίδρασης με το μοντέλο

Α ομάδα (κορίτσι-αγόρι)

Αρχικά, η ομάδα ανέλυσε προφορικά την 1η εικόνα αναφοράς, ζήτησε στυλό και χαρτί για να γράψει την περιγραφή της και μετά επέλεξε να υπαγορεύσει την προτροπή στα ελληνικά. Το μοντέλο δημιούργησε τέσσερις εναλλακτικές εικόνες, αλλά καμία δεν θεωρήθηκε από τα παιδιά ότι αντιστοιχούσε στην εικόνα αναφοράς. Στη συνέχεια, προσπάθησαν με νέα υπαγόρευση να διορθώσουν την αρχική προτροπή, όμως το μοντέλο σταμάτησε να ανταποκρίνεται, αδυνατώντας να παραγάγει τη ζητούμενη εικόνα. Στην τελευταία προσπάθεια των παιδιών για επανεκφώνηση της προτροπής, το μοντέλο έκανε μετεγγραφή

της ελληνικής εκφώνησης σε λατινικούς χαρακτήρες και σταμάτησε να ανταποκρίνεται. Συνεπώς, η ομάδα κατέληξε στην επιλογή μιας εικόνας από την αρχική προτροπή, που έμοιαζε περισσότερο με την εικόνα αναφοράς. Στη 2η εικόνα αναφοράς, τα παιδιά άλλαξαν τακτική αποφασίζοντας να γράψουν την προτροπή στα ελληνικά και να μην την υπαγορεύσουν. Το μοντέλο ανταποκρίθηκε παράγοντας τέσσερις εικόνες, οι οποίες όμως είχαν διαφορετικό στυλ από αυτό της εικόνας αναφοράς. Στη προοπτική τους να το διορθώσουν, τα παιδιά επέλεξαν να εκφωνήσουν αυτή τη φορά στα αγγλικά τη νέα προτροπή. Το μοντέλο ανταποκρίθηκε και παρήγαγε εικόνες που, πάλι, δεν ήταν συμβατές με το στυλ της εικόνας αναφοράς, αλλά επιλέχθηκε η εικόνα που έμοιαζε περισσότερο. Κατά την αναπαραγωγή της 3ης εικόνας αναφοράς, η ομάδα Α επέλεξε να γράψει την προτροπή στα ελληνικά, εφόσον ήταν μικρή σε έκταση, αλλά το μοντέλο δεν ανταποκρίθηκε. Επιχείρησαν δεύτερη φορά να υπαγορεύσουν την προτροπή, αλλά τα αποτελέσματα των εικόνων που δημιουργήθηκαν δεν τους ικανοποίησαν. Προσπάθησαν εκ νέου να υπαγορεύσουν την προτροπή αργά και δυνατά στα αγγλικά, αλλά οι εικόνες που παρήγαγε το μοντέλο δεν ήταν ικανοποιητικές για τα παιδιά, τα οποία τελικά επέλεξαν την πιο σχετική εικόνα από προηγούμενη προτροπή. Σε όλες τις επόμενες εικόνες αναφοράς η ομάδα ακολούθησε τη διαδικασία υπαγόρευσης των προτροπών στα ελληνικά.

Β ομάδα (κορίτσια)

Η συγκεκριμένη ομάδα ξεκίνησε την αναπαραγωγή της 1ης εικόνας αναφοράς γράφοντας τις προτροπές στα ελληνικά. Βλέποντας τις παραγόμενες εικόνες και μετά από συνεννόηση τα κορίτσια θέλησαν να κάνουν μια τροποποίηση της προτροπής και έγραψαν στο μοντέλο: «*Θέλω να προσθέσω κάτι ακόμη στην περιγραφή*». Το μοντέλο απάντησε: «*Φυσικά, πες μου τι θα ήθελες να προσθέσεις και θα το ενσωματώσω στην περιγραφή της εικόνας*». Το μοντέλο ανταποκρίθηκε στην προσθήκη και δημιούργησε αντίστοιχες εικόνες. Τα παιδιά επέλεξαν την εικόνα που έμοιαζε περισσότερο με την 1η εικόνα αναφοράς. Σε όλες τις επόμενες εικόνες αναφοράς η ομάδα ακολούθησε το ίδιο μοτίβο αλληλεπίδρασης με το μοντέλο.

Γ ομάδα (αγόρια)

Η ομάδα αυτή ξεκίνησε την αναπαραγωγή της 1ης εικόνας αναφοράς γράφοντας την προτροπή στα ελληνικά. Το μοντέλο ανταποκρίθηκε στην προτροπή και τα παιδιά επέλεξαν την σχετική εικόνα. Η ίδια διαδικασία αλληλεπίδρασης και αναθεώρησης προτροπών επαναλήφθηκε για την αναπαραγωγή των υπόλοιπων εικόνων αναφοράς, εκτός της 9ης εικόνας όπου η προτροπή γράφτηκε στα αγγλικά.

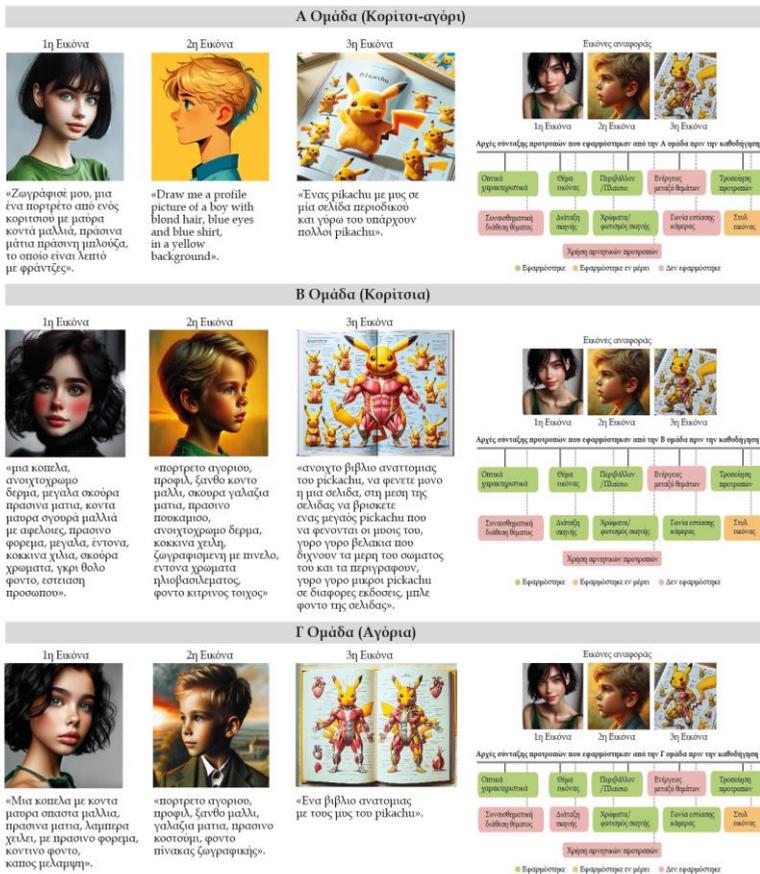
Από την παραπάνω αλληλεπίδραση των ομάδων με τη διεπαφή του μοντέλου, αναδεικνύεται η σημασία του τρόπου αλληλεπίδρασής τους με αυτό, για την επίτευξη του επιθυμητού αποτελέσματος. Η διεπαφή του μοντέλου παρέχει τη δυνατότητα πολυτροπικής αλληλεπίδρασης για την είσοδο των προτροπών, η οποία ενέχει πλεονεκτήματα και μειονεκτήματα. Τα πλεονεκτήματα έγκεινται στην ευχρηστία, την ευελιξία και την προσβασιμότητα του συστήματος μέσω της φυσικής γλώσσας. Τα μειονεκτήματα αφορούν τη δυσκολία δημιουργίας μιας επιδιωκόμενης εικόνας με την επιλογή εκφώνησης των προτροπών στα ελληνικά (καθώς μειώνεται η απόδοση των μοντέλων κατά τη χρήση άλλων γλωσσών εκτός της αγγλικής) (Reviriego & Merino-Gómez, 2022) και την εμφάνιση τεχνικών περιορισμών («πάγωμα») του μοντέλου, ιδίως κατά τη διαδικασία εναλλαγής της τροπικότητας.

Θέμα: Αυτόνομη ανάπτυξη στρατηγικών εφαρμογής προτροπών

Όλες οι ομάδες ανέπτυξαν στρατηγικές προτροπών πριν την σταδιακή παρουσίασή τους, από την πρώτη συγγραφέα. Σε αυτό εκτιμάται ότι οδηγήθηκαν μέσα από την ομαδική συνεργασία, τον καταιγιισμό ιδεών και τον πειραματισμό (διαδικασία δοκιμής-λάθους-επανασχεδιασμού προτροπών). Συγκεκριμένα, στον πίνακα 1 που δείχνει τις προτροπές κάθε ομάδας για τη δημιουργία των τριών πρώτων εικόνων αναφοράς, φαίνεται ότι όλες οι ομάδες αναγνώρισαν τα οπτικά χαρακτηριστικά, επέλεξαν λέξεις-κλειδιά για την περιγραφή τους, εντόπισαν τα χρώματα, τα θέματα (αντικείμενα-χαρακτήρες) και περιέγραψαν τα χαρακτηριστικά τους. Επανασχεδίασαν τις κειμενικές προτροπές τους για να δημιουργήσουν τις επιθυμητές εικόνες, περιέγραψαν τον φωτισμό της σκηνής και προσδιόρισαν το περιβάλλον κατά τη δημιουργία της 3ης εικόνας.

Η Α ομάδα στη 2η και 3η εικόνα επισήμανε τις σχέσεις διάταξης σκηνής μεταξύ των υποκειμένων. Το ίδιο έκανε και η Β ομάδα στην 3η εικόνα. Επιπλέον, η Β και Γ ομάδα στην 1η εικόνα ανέφεραν την εστίαση της κάμερας. Τέλος, στη 2η εικόνα αναφοράς που αναφέρεται σε πορταίτο αγοριού σε στυλ ελαιογραφίας, όλες οι ομάδες εντόπισαν μόνο το πορταίτο και όχι το στυλ.

Πίνακας 1. Εφαρμογή αρχών για την αναπαραγωγή εικόνων χωρίς καθοδήγηση



Θέμα: Παροχή και ενσωμάτωση στρατηγικών αρχών προτροπών

Στις τρεις ομάδες, κατά την αναπαραγωγή των τεσσάρων εικόνων αναφοράς που ακολούθησαν, δόθηκαν σταδιακά από την πρώτη συγγραφέα οι έντεκα αρχές προτροπών, παρά το γεγονός ότι κάποιες από αυτές (περίπου οι μισές) είχαν ήδη αναδυθεί και χρησιμοποιηθεί από τα παιδιά πριν την καθοδήγηση.

Α ομάδα (κορίτσι-αγόρι)

Στο πλαίσιο αυτό, για την αναπαραγωγή της 4ης εικόνας αναφοράς η Α ομάδα υπαγόρευσε στο μοντέλο την εξής προτροπή: *«Πολλοί Σπαρτιάτες γιορτάζουν σε παραλία και πάνω από πάνω τους υπάρχουν αετοί και ο ουρανός είναι συννεφιασμένος»*, ενσωματώνοντας στην προτροπή τη λέξη κλειδί *«Σπαρτιάτες»*. Στην παραπάνω προτροπή εντοπίζεται επίσης και η εισαγωγή της συναισθηματικής διάθεσης του θέματος, που εφαρμόστηκε από την ομάδα, χωρίς να έχει ακόμα δοθεί ως αρχή προτροπής. Αντίστοιχα, το ίδιο συνέβη και με την αρχή περιγραφής των ενεργειών μεταξύ των χαρακτήρων της εικόνας αναφοράς.

Κατά την αναπαραγωγή της 5ης εικόνας αναφοράς η συγκεκριμένη ομάδα εκφώνησε την παρακάτω προτροπή: *«Δημιούργησε μου μια καρτουίστικη εικόνα με μια γάτα αστροναύτη που κάθεται σε μια ξαπλώστρα στον πλανήτη Άρη»*. Στην προτροπή αυτή εντοπίζεται η ρητή αναφορά *«καρτουίστικη εικόνα»* και εδραιώνεται η εφαρμογή της αρχής προτροπής που αφορά το στυλ της εικόνας. Φαίνεται ακόμα ότι, η αρχή της περιγραφής των ενεργειών, χρησιμοποιήθηκε και ενσωματώθηκε στις στρατηγικές των παιδιών. Στην αναπαραγωγή της 6ης εικόνας αναφοράς, παρατηρήθηκε ότι τα παιδιά εδραιώσαν στις προτροπές τους τις αρχές διάταξης σκηνής και χρωμάτων.

Τα παιδιά, για την αναπαραγωγή της 7ης εικόνας αναφοράς, υπαγόρευσαν την εξής προτροπή: *«Ένα μπολ σε ένα τραπέζι με πατάτες στα δεξιά, φράουλες στα αριστερά, στη μέση κρέας και βασιλικό»*. Πριν τη δημιουργία της 7ης εικόνας αναφοράς, δόθηκαν στην Α ομάδα οι τελευταίες αρχές προτροπών που ήταν η γωνία εστίασης κάμερας/πλάνο, το στυλ της εικόνας και η χρήση των αρνητικών προτροπών. Η γωνία εστίασης και η χρήση της αρνητικής προτροπής δε χρησιμοποιήθηκε τη δεδομένη στιγμή από τη συγκεκριμένη ομάδα.

Β ομάδα (κορίτσια)

Η ομάδα αυτή κατά τη διάρκεια αναπαραγωγής της 4ης εικόνας αναφοράς, έγραψε την προτροπή: *«μια εικόνα που απεικονίζει τη συμφιλίωση δύο ομάδων σπαρτιατων πολεμιστών. Πέντε αρχηγοί πολεμιστές βρίσκονται μπροστά, οι τρεις από αυτούς φορούν πανοπλία και κράνος, ενώ οι άλλοι δύο αρχηγοί του αντίπαλου στρατοπέδου φορούν μόνο κράνος. Κάποιοι κρατούν ποτήρια με κρασί και άλλοι διάφορα δώρα. Βρίσκονται σε μια παραλία και περπατούν δίπλα στη θάλασσα, ενώ πίσω τους ακολουθούν πολλοί στρατιώτες. Ο καιρός είναι συννεφιασμένος αλλά έχει εντονα χρωματα, μεγάλα χρωματιστά μπαλόνια βρίσκονται στον ουρανό, στο έδαφος υπάρχουν πεταμένα μπουκάλια, και στον αέρα πετούν γεράκια»*. Στην παραπάνω προτροπή παρατηρείται αρχικά μία εκτενής περιγραφή της εικόνας αναφοράς και ότι δεν επιχειρήθηκε από την ομάδα επανασχεδιασμός της, με σκοπό τη συγχώνευση λέξεων-κλειδιών, για τη δημιουργία πιο συνοπτικών προτροπών, πιθανώς γιατί ικανοποιήθηκαν από το αποτέλεσμα της εικόνας. Ακόμα, παρατηρείται ότι η συγκεκριμένη ομάδα εντόπισε και εφάρμοσε αυτόνομα δύο νέες αρχές καθοδήγησης προτροπών, τη συναισθηματική διάθεση του θέματος και την αρχή περιγραφής των ενεργειών μεταξύ των χαρακτήρων. Τα παιδιά για την αναπαραγωγή της 5ης εικόνας αναφοράς ενσωμάτωσαν στην προτροπή την αρχή της περιγραφής των ενεργειών, αλλά δεν επιχειρήσαν τον επανασχεδιασμό της προτροπής.

Στην αναπαραγωγή της 6ης εικόνας αναφοράς, η Β ομάδα έγραψε την προτροπή: *«ενα ξυλινο πειρατικο караβι στην ερημο, γυρο απο το караβι να υπαρχει ομηχλη στις αποχρωσεις του ροζ*

και του μοβ, βροχη μετεωρητων, αστερια στον ουρανο, αριστερα πανω να φαινεται το φεγγαρι, μαγικο τοπιο, ροζ και μοβ συννεφα στον ουρανο, πισω απο το караβι να φαινεται ηλιοβασιλεμα». Στην παραπάνω προτροπή παρατηρείται η πλήρης ενσωμάτωση και υιοθέτηση των αρχών διάταξης σκηνής και των χρωμάτων. Αντίστοιχα το ίδιο συνέβη και στην προτροπή που έγραψε η ομάδα για την αναπαραγωγή της 7ης εικόνας αναφοράς, εκτός από την αρχή της αρνητικής προτροπής που ενώ δώθηκε δεν χρησιμοποιήθηκε από την ομάδα.

Γ ομάδα (αγόρια)

Η συγκεκριμένη ομάδα για την αναπαραγωγή της 4ης εικόνας αναφοράς, έγραψε την προτροπή: «Χιλιάδες δυνατοι και μυοδες σπαρτιάτες που είναι στον δρομο για να γυρισουν πισω στην πατριδα τους γιορταζοντας μια νικη τους δηπλα στην ακτή πινοντας κρασι με συνεφιασμενο ουρανο και με πολυχρομες μπαλες στον αερα». Σε αυτή την προτροπή εντοπίστηκε και ενσωματώθηκε η λέξη κλειδί «Σπαρτιάτες», ενώ παρατηρήθηκε ότι τα παιδιά εφάρμοσαν από μόνα τους (χωρίς να έχει δοθεί) δύο νέες αρχές καθοδήγησης προτροπών, της συναισθηματικής διάθεσης του θέματος και της περιγραφής ενεργειών μεταξύ των χαρακτήρων, σημειώνοντας πρόοδο.

Στην αναπαραγωγή της 5ης εικόνας αναφοράς η ομάδα έγραψε την προτροπή: «Μια γατα με ενα πρασινο και ενα μπλε ματι οπου το κεφαλι της απο την αριστερη πλευρα είναι πρασινο και απο την δεξια πλευρα είναι πορτοκαλι ενω στο κεντρο είναι ασπρη, φοραει στολη αστροναυτη και καθετε σε μια πλαστικη ξαπλωστρα, εχοντας διπλα της ενα μικρο μπλε μπουκαλι, στον πλανητη Αρη». Στη προτροπή αυτή αναφέρεται ρητά και εφαρμόζεται η αρχή διάταξη σκηνής, χωρίς να έχει ακόμα δοθεί. Κατά την αναπαραγωγή της 6ης εικόνας αναφοράς, παρατηρήθηκε ότι τα παιδιά εδραίωσαν στις προτροπές τους τις αρχές διάταξης σκηνής και χρωμάτων.

Στο πλαίσιο αναπαραγωγής της 7ης εικόνας αναφοράς η Γ ομάδα έγραψε την προτροπή: «πιάτο με κρέας και πατάτες, φράουλες, blueberrys, βασιλικό και δύο μικρα μπολακια με σάλτσα και μαγιονεζα, όλα σε ρεαλιστικό και πολύ κοντινό και πλαγιο πλάνο, όχι μπολ μέσα στο πιάτο». Στο σημείο αυτό εντοπίζεται η εδραίωση της αρχής γωνίας εστίασης κάμερας/πλάνου, η ρητή αναφορά στο στυλ της εικόνας και η υιοθέτηση της αρχής των αρνητικών προτροπών. Η Γ ομάδα, ολοκληρώνοντας τις τέσσερις εικόνες, εφάρμοσε και ενσωμάτωσε όλες τις παρεχόμενες αρχές προτροπών.

Συνολικά, από την παραπάνω ανάλυση προκύπτει ότι οι συμμετέχοντες/ουσες ανέπτυξαν και ενσωμάτωσαν μια σειρά στρατηγικών προτροπών, με τις οποίες μπορούσαν να δημιουργήσουν τις επιθυμητές εικόνες. Τα παιδιά ήταν δημιουργικά και ενσωμάτωσαν πλούσια περιγραφική γλώσσα στις προτροπές τους, χωρίς όμως να έχουν χρησιμοποιήσει συγκεκριμένες λέξεις-κλειδιά. Κατά συνέπεια, οι προτροπές τους ήταν εκτεταμένες έχοντας περισσότερες πιθανότητες να περιλαμβάνουν κατά τύχη λέξεις-κλειδιά, οι οποίες θα μπορούσαν να ενεργοποιήσουν το μοντέλο για τη σύνθεση της επιθυμητής εικόνας (Oppenlaender et al., 2023). Όσον αφορά τη μη χρήση της αρνητικής προτροπής, ορισμένοι συμμετέχοντες ίσως να μην κατάλαβαν την έννοια του όρου, παρόλο που εξηγήθηκε. Συμπερασματικά, η χρήση της φυσικής γλώσσας μπορεί να είναι ιδιαίτερα διαισθητική και βοηθητική για την αλληλεπίδραση με το μοντέλο, αλλά δεν συμβαίνει το ίδιο με τη διαδικασία αναζήτησης βέλτιστης προτροπής (prompt engineering), μια δεξιότητα που δεν είναι διαισθητική και πρέπει να αποκτηθεί μέσω μάθησης και εξάσκησης, ώστε να συμβάλλει στη δημιουργία αποτελεσματικών προτροπών (Oppenlaender et al., 2023).

Θέμα: Έλεγχος αυτόνομης ενσωμάτωσης παρεχόμενων αρχών προτροπών

Σε αυτό το στάδιο οι ομάδες κλήθηκαν να αναπαραγάγουν αυτόνομα δύο ακόμη εικόνες αναφοράς με σκοπό τον έλεγχο του βαθμού ενσωμάτωσης των αρχών για τη δημιουργία προτροπών, μετά την καθοδήγηση.

A ομάδα (κορίτσι-αγόρι)

Η ομάδα αυτή κατά την αναπαραγωγή της 8ης εικόνας αναφοράς, υπαγόρευσε στο μοντέλο την προτροπή: «Μια φωτογραφία μινιατούρας μουσικών που παίζουν μουσική σε μια πλατεία γύρω από μαγαζιά». Ενώ, για την αναπαραγωγή της 9ης εικόνας αναφοράς υπαγόρευσε την προτροπή: «Ένα υπνοδωμάτιο με ένα μονό κρεβάτι που πάνω του υπάρχουν μωβ μπαλόνια και δεξιά, μια αφίσα του Χάρι Πότερ». Από τις παραπάνω προτροπές διαπιστώνεται ότι η Α ομάδα, εστίασε στα κύρια οπτικά χαρακτηριστικά των εικόνων αναφοράς και έκανε συγκεκριμένη επιλογή λέξεων-κλειδιών για την περιγραφή των εικόνων, χωρίς λεπτομέρειες και μακροσκελείς περιγραφές. Ωστόσο, η γωνία εστίασης/πλάνο και η χρήση της αρνητικής προτροπής δεν χρησιμοποιήθηκαν από την Α ομάδα.

B ομάδα (κορίτσια)

Η συγκεκριμένη ομάδα για την αναπαραγωγή της 8ης εικόνας αναφοράς, έγραψε την προτροπή: «μια μαकेτα αυλής εστιατοριο με μινιατουρες που παιζουνε μουσικη, πετρινο πατωμα, δυο κτηρια που σχηματιζουν μια γωνια με πολλα λουλουδια, το ενα κτηριο εχει ριγε τεντες με κοκκινο και ασπρο, καποιες μινιατουρες καθοντε σε τραπεζια και παρακολουθουν ενω καποιες αλλες παιζουν μουσικη και κρατουν οργανα, εντανα χρωματα, εστειασει στους μουσικους και εχει λιγο θολο φοντο». Ενώ για την αναπαραγωγή της 9ης εικόνας αναφοράς, έγραψε την προτροπή: «ενα δωματιο με ενα μωβ κρεβατι στη μεση, να ειναι γωνιακο το δωματιο, ενα πατακι στις αποχρωσεις του μωβ, πισο απο το κρεβατι εχει μια ταπετσορια στον τοιχο που εχει χαρακτηρισες του χαρι ποτερ, στον αλλο τοιχο να εχει μια μεγαλη αφισα του χαρι ποτερ, στην γωνια εκει που χωριζοντα οι δυο τοιχοι εχει μια μικρη βιβλιοθηκη, πανω απο το κρεβατι εχει ενα φωτηστικο που κρεμετε απο πολλα μπλε μωβ και ασπρα μπαλονια, αριστερα και δεξια απο το κρεβατι εχει απο ενα κυλινδρικο ξυλινο κομοδινο, το πατωμα ειναι ξυλινο». Από τις παραπάνω προτροπές διαπιστώνεται ότι η Β ομάδα εξακολουθούσε να διατηρεί την αναλυτική λεπτομερή περιγραφική της πορεία, χωρίς να έχει εστιάσει στα κύρια οπτικά χαρακτηριστικά ώστε να επιλέγει συγκεκριμένες λέξεις-κλειδιά, που θα οδηγούσαν σε μία περιγραφή μικρότερης έκτασης. Ωστόσο, έγινε ρητή αναφορά στο στυλ και επισημάνθηκε η εστίαση της κάμερας στους μουσικούς, ενώ δεν έγινε χρήση της αρνητικής προτροπής.

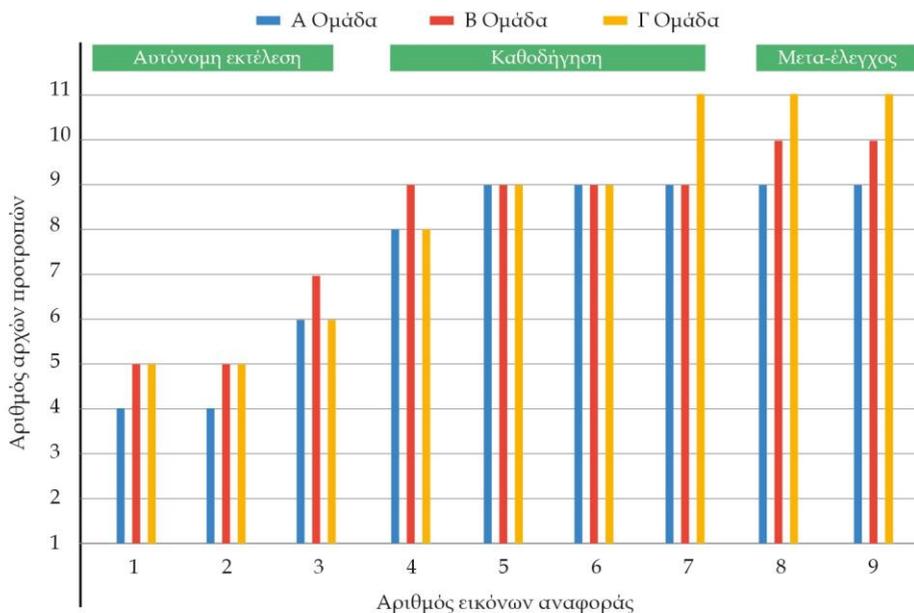
Γ ομάδα (αγόρια)

Αυτή η ομάδα για την αναπαραγωγή της 8ης εικόνας αναφοράς, έγραψε την προτροπή: «μια 3D μινιατούρα, καφε, βράδυ, μπαντα που παίζει μουσική, πέτρινη πλατεία, κοσμο, κοντινό πλάνο, λήψη απο ψηλά και ζεστά χρώματα». Ενώ για την αναπαραγωγή της 9ης εικόνας αναφοράς, έγραψε την προτροπή για πρώτη φορά στα αγγλικά (μετά από «πάγωμα» του μοντέλου): «*bedroom, purple bed frame, harry potter poster and purple and white ballons, lamp on top of the bed with baloons!*». Από τις παραπάνω προτροπές διαφαίνεται ότι η Γ ομάδα, εστίασε στα κύρια οπτικά χαρακτηριστικά των εικόνων αναφοράς και έκανε συγκεκριμένη επιλογή λέξεων-κλειδιών, για την περιγραφή των εικόνων, χωρίς λεπτομέρειες και μακροσκελείς περιγραφές. Το στυλ όσο και η γωνία εστίασης/πλάνο χρησιμοποιήθηκαν αποτελεσματικά από την ομάδα, καθώς φαίνεται ότι ενσωμάτωσε τις αρχές και βελτίωσε τις προτροπές της.

Συμπερασματικά οι αρχές προτροπών που χρησιμοποίησαν περισσότερο οι ομάδες ήταν: Η αναγνώριση των οπτικών χαρακτηριστικών, οι λέξεις κλειδιά, ο εντοπισμός του θέματος, ο προσδιορισμός του περιβάλλοντος, η περιγραφή των ενεργειών, η περιγραφή της συναισθηματικής διάθεσης του θέματος, ο επανασχεδιασμός των κειμενικών προτροπών, η περιγραφή των χρωμάτων και του φωτισμού της σκηνής και η περιγραφή των σχέσεων διάταξης σκηνής. Η αρχή αναφοράς της γωνίας εστίασης κάμερας/πλάνο ήταν αυτή που χρησιμοποιήθηκε από τη Β και Γ ομάδα, στην 8η και 9η εικόνα, ενώ δεν χρησιμοποιήθηκε από την ομάδα Α. Οι αρνητικές προτροπές ενσωματώθηκαν μόνο από την Γ ομάδα στις

παραπάνω εικόνες. Η αναπαραγωγή της 8ης εικόνας δυσκόλεψε περισσότερο όλες τις ομάδες, λόγω της ιδιαιτερότητας του στυλ (φωτογραφία μινιατούρας) και της γωνίας λήψης της κάμερας (ψηλή γωνία). Στο Γράφημα 1 παρουσιάζεται η συνολική πορεία εφαρμογής του αριθμού των αρχών προτροπών κάθε ομάδας, ανά εικόνα αναφοράς.

Γράφημα 1. Αριθμός αρχών προτροπών που εφαρμόστηκαν ανά ομάδα, ανά εικόνα



Θέμα: Αξιολόγηση αλληλεπίδρασης με το μοντέλο

Αναφορικά με την αξιολόγηση της διαδικασίας αλληλεπίδρασης με το συνθετικό μέσο DALL-E, η δημιουργία συνθετικών εικόνων χαρακτηρίστηκε ως ενδιαφέρουσα από τα παιδιά της Α και Β ομάδας και ως μέτρια από αυτά της Γ ομάδας. Οι ομάδες Α και Β δήλωσαν ότι ικανοποιήθηκαν από τις περισσότερες συνθετικές εικόνες που δημιούργησαν και τις χαρακτήρισαν παρόμοιες με τις εικόνες αναφοράς. Αντίθετα, η Γ ομάδα δήλωσε ότι οι εικόνες που παρήγαγε το μοντέλο δεν ήταν ικανοποιητικές, γιατί οι θέσεις συγκεκριμένων αντικειμένων, χρωμάτων και υλικών ήταν διαφορετικές και το μοντέλο έδειχνε να αγνοεί τις σχετικές προτροπές τους. Οι αρχικές σκέψεις των παιδιών της Γ ομάδας ήταν ότι θα ήταν εύκολο και θα περνούσαν ωραία αλλά: «...στην τελική έγινε το αντίθετο. Και ήταν δύσκολο. Ήταν δύσκολο και κουρασθήκαμε πολύ». Σε καμία ομάδα δεν άρεσε το γεγονός ότι πολλές φορές το μοντέλο σταματούσε να ανταποκρίνεται και ήταν ένας από τους κύριους λόγους που κουράστηκαν και απογοητεύτηκαν. Όλα τα παιδιά εξέφρασαν την επιθυμία, στο άμεσο μέλλον, να εξελιχθούν τα παραγωγικά μοντέλα τεχνητής νοημοσύνης, ώστε να υπάρχει καλύτερη αλληλεπίδραση, επικοινωνία και να παρέχει τα επιθυμητά αποτελέσματα.

Συζήτηση

Στη μελέτη αυτή εξετάστηκαν οι πορείες δημιουργίας προτροπών (prompt engineering) τριών ομάδων παιδιών ηλικίας 12-14 ετών, για τη δημιουργία συνθετικών εικόνων με τη χρήση της παραγωγικής τεχνητής νοημοσύνης. Παρόλο που οι συμμετέχοντες ενεπλάκησαν για πρώτη φορά συστηματικά με το συνθετικό μέσο τεχνητής νοημοσύνης DALL-E, με σκοπό τη δημιουργία εικόνων, επέδειξαν ανοιχτό πνεύμα στην ιδέα και προσαρμόστηκαν γρήγορα στη χρήση του. Καθ' όλη τη διάρκεια της παρέμβασης, τα παιδιά βίωσαν μια συναισθηματική τροχιά ξεκινώντας με αρχικό ενδιαφέρον, ενώ ακολούθησε μια περίοδος ταχείας εξερεύνησης και προθυμίας για τη δημιουργία συνθετικών εικόνων. Στην πορεία όμως βίωσαν απογοήτευση, γιατί αντιμετώπισαν τους περιορισμούς του συνθετικού μέσου σε σχέση με τις παρεχόμενες προτροπές και τις παραγόμενες εικόνες, ευρήματα που συμφωνούν με τους Hutson και Lang (2023). Ωστόσο, οι περιορισμοί αυτοί οδήγησαν τα παιδιά στον πειραματισμό και την ανάπτυξη εναλλακτικών προσεγγίσεων και στρατηγικών. Παρόλο που αρχικά οι εικόνες που δημιουργήσαν τα παιδιά έμοιαζαν συναρπαστικές, σύντομα αντιλήφθηκαν ότι το μοντέλο περιόριζε τις δημιουργικές τους δυνατότητες. Αυτό σημαίνει ότι τα εργαλεία παραγωγικής τεχνητής νοημοσύνης μπορούν να χρησιμοποιηθούν ως πολύτιμες πηγές έμπνευσης και καθοδήγησης για τα παιδιά, αλλά δεν μπορούν να υποκαταστήσουν τη δημιουργικότητά τους (Hutson & Lang, 2023).

Στη μελέτη επισημάνθηκε από τα παιδιά ότι το μοντέλο έδειχνε να αγνοεί κάποιες χαρακτηριστικές λεπτομέρειες στις σχετικές προτροπές τους, χωρίς να δημιουργεί τις επιθυμητές εικόνες. Αυτό συνέβαινε γιατί στην πράξη οι λέξεις-κλειδιά που χρησιμοποιούν οι άνθρωποι για να περιγράψουν ένα θέμα μπορεί να μην αντιστοιχούν στις λέξεις που έχει μάθει το νευρωνικό δίκτυο, κατά τη διάρκεια της εκπαίδευσής του, με αποτέλεσμα το ίδιο το μοντέλο να πρέπει να συμπληρώσει τα κενά που δημιουργούνται (Oppenlaender, 2022). Με τη σειρά του, το γεγονός αυτό μπορεί να οδηγήσει τον χρήστη είτε να συμβιβαστεί με μια εικόνα που είναι «αρκετά καλή» είτε να εγκαταλείψει την αναζήτηση της αρχικά οραματιζόμενης εικόνας (Oppenlaender, 2022), κάτι το οποίο παρατηρήθηκε και στην παρούσα μελέτη. Από την άλλη πλευρά, η αναντιστοιχία μεταξύ του επιθυμητού αποτελέσματος και της παραγόμενης εικόνας μπορεί να πυροδοτήσει τη δημιουργικότητα του χρήστη (Erstein et al., 2022). Εφόσον ο τομέας της γενετικής τεχνητής νοημοσύνης συνεχίζει να εξελίσσεται με ταχείς ρυθμούς, αυτοί οι περιορισμοί είναι πιθανό να αντιμετωπιστούν, καθιστώντας τα συνθετικά εργαλεία πιο προσιτά και αποτελεσματικά για την επικοινωνία και την αλληλεπίδραση με τους χρήστες (Hutson & Lang, 2023· Oppenlaender, 2023). Τέλος, πρέπει να επισημανθεί ότι οποιαδήποτε άλλη ακολουθία ή σύνθεση εικόνων θα μπορούσε να οδηγήσει στη συλλογή διαφορετικών δεδομένων και αυτό συγκαταλέγεται στους περιορισμούς της μελέτης.

Βιβλιογραφικές Αναφορές

- Ali, S., Ravi, P., Williams, R., DiPaola, D., & Breazeal, C. (2024). Constructing dreams using generative AI. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 21, pp. 23268-23275). <https://doi.org/10.1609/aaai.v38i21.30374>
- Atlas, S. (2023). ChatGPT for higher education and professional development: A guide to conversational AI. https://digitalcommons.uri.edu/cba_facpubs/548
- Audry, S. (2021). *Art in the age of machine learning*. MIT Press.
- Baidoo-Anu, D., & Anshah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62.

- Caramiaux, B., & Fdili Alaoui, S. (2022). "Explorers of Unknown Planets" Practices and Politics of Artificial Intelligence in Visual Arts. In the *Proceedings of the ACM on Human-Computer Interaction* (pp. 1-24).
- Choi, D., Hong, S., Park, J., Chung, J. J. Y., & Kim, J. (2023). CreativeConnect: Supporting Reference Recombination for Graphic Design Ideation with Generative AI. *arXiv preprint arXiv:2312.11949*. <https://doi.org/10.48550/arXiv.2312.11949>
- Dodić, D., & Čungurski, S. (2023). The Picture World of the Future: AI Text-to-image as a New Era of Visual Content Creation. *KNOWLEDGE-International Journal*, 57(3), 417-421. Retrieved from <http://ikm.mk/ojs/index.php/kij/article/view/6010>
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing*, 62(1), 107-115.
- Epstein, Z., Schroeder, H., & Newman, D. (2022). When happy accidents spark creativity: Bringing collaborative speculation to life with generative AI. *arXiv preprint arXiv:2206.00533*.
- Fiebrink, R. (2019). Machine learning education for artists, musicians, and other creative practitioners. *ACM Transactions on Computing Education (TOCE)*, 19(4), 1-32.
- Gu, J., Han, Z., Chen, S., Beirami, A., He, B., Zhang, G., ... & Torr, P. (2023). A systematic survey of prompt engineering on vision-language foundation models. *arXiv preprint arXiv:2307.12980*. <https://doi.org/10.48550/arXiv.2307.12980>
- Hutson, J., & Lang, M. (2023). Content creation or interpolation: AI generative digital art in the classroom. *Metaverse*, 4(1), 13. <https://doi.org/10.54517/m.v4i1.2158>
- Hwang, Y., Lee, J. H., & Shin, D. (2023). What is prompt literacy? An exploratory study of language learners' development of new literacy skill using generative AI. *arXiv preprint arXiv:2311.05373*. <https://doi.org/10.48550/arXiv.2311.05373>
- Kim, S., Ko, T., Kwon, Y., & Lee, K. (2023). Designing interfaces for text-to-image prompt engineering using stable diffusion models: a human-AI interaction approach. <https://doi.org/10.21606/iasdr.2023.448>
- Manovich, L., Manovich, L., & Arielli, E. (2023). AI Image and Generative Media. *Artificial Aesthetics: A Critical Guide to AI, Media and Design*. Retrieved April 18, 2024, from <http://manovich.net/index.php/projects/artificial-aesthetics-book>
- Microsoft. (2024). *Microsoft Copilot Pro*. Retrieved May 3, 2024, from <https://copilot.microsoft.com>
- Oppenlaender, J. (2023). The Cultivated Practices of Text-to-Image Generation. *arXiv preprint arXiv:2306.11393*. <https://doi.org/10.48550/arXiv.2306.11393>
- Oppenlaender, J., Linder, R., & Silvennoinen, J. (2023). Prompting ai art: An investigation into the creative skill of prompt engineering. *arXiv preprint arXiv:2303.13534*. <https://doi.org/10.48550/arXiv.2303.13534>
- Oppenlaender, J. (2022). Prompt engineering for text-based generative art. *arXiv preprint arXiv:2204.13988*, 2.
- Reviriego, P., & Merino-Gómez, E. (2022). Text to image generation: Leaving no language behind. *arXiv preprint arXiv:2208.09333*. <https://doi.org/10.48550/arXiv.2208.09333>
- Sanchez, T. (2023). Examining the Text-to-Image Community of Practice: Why and How do People Prompt Generative AIs?. In *Proceedings of the 15th Conference on Creativity and Cognition* (pp. 43-61). <https://doi.org/10.1145/3591196.3593051>
- Taheri, A., Izadi, M., Shriram, G., Rostamzadeh, N., & Kane, S. (2023). Breaking Barriers to Creative Expression: Co-Designing and Implementing an Accessible Text-to-Image Interface. *arXiv preprint arXiv:2309.02402*. <https://doi.org/10.48550/arXiv.2309.02402>
- Vartiainen, H., & Tedre, M. (2024). How Text-to-Image Generative AI is Transforming Mediated Action. *IEEE Computer Graphics and Applications*. <https://doi.org/10.1109/MCG.2024.3355808>
- Vartiainen, H., Tedre, M., & Jormanainen, I. (2023). Co-creating digital art with generative AI in K-9 education: Socio-material insights. *International Journal of education through art*, 19(3), 405-423. https://doi.org/10.1386/eta_00143_1
- Venkataramanan, A. (2023). *stability. gpt: Combining GPT-4 and Stable Diffusion to Generate Storybooks That Are Textually and Visually Cohesive*. University of California, Santa Cruz. <https://www.proquest.com/openview/72d8cbb22040970f665f2650913e5cc7/1?pq-origsite=gscholar&cbl=18750&diss=y>

Woo, D. J., Guo, K., & Susanto, H. (2023). Cases of EFL secondary students' prompt engineering pathways to complete a writing task with ChatGPT. *arXiv preprint arXiv:2307.05493*. <https://doi.org/10.48550/arXiv.2307.05493>