



Η ΜΟΥΣΙΚΗ ΣΤΗΝ ΨΗΦΙΑΚΗ ΕΠΟΧΗ

ΔΙΕΘΝΕΣ ΣΥΝΕΔΡΙΟ | ΨΗΦΙΑΚΕΣ ΤΕΧΝΟΛΟΓΙΕΣ & ΔΙΚΑΙΩΜΑΤΑ ΤΩΝ ΚΑΛΛΙΤΕΧΝΩΝ

ΑΘΗΝΑ, ΟΚΤΩΒΡΙΟΣ 22-24, ΚΕΝΤΡΟ ΠΟΛΙΤΙΣΜΟΥ ΙΔΡΥΜΑ ΣΤΑΥΡΟΣ ΝΙΑΡΧΟΣ

ΠΑΡΟΥΣΙΑΣΕΙΣ

“Η Μουσική Στην Ψηφιακή Εποχή: Streaming & Τεχνητή Νοημοσύνη”, ήταν ένα διεθνές forum που συνδιοργάνωσε ο ΑΠΟΛΛΩΝ (ΟΣΔ συγγενικών δικαιωμάτων Ελλήνων μουσικών) και η FIM (Διεθνής Ομοσπονδία Μουσικών). Καθώς το **περιεχόμενο που παράγεται από Τεχνητή Νοημοσύνη και η κυριαρχία των πλατφορμών streaming εξαπλώνονται**, το συνέδριο εξέτασε πώς **τα μοντέλα εσόδων, τα νομικά πλαίσια και η καλλιτεχνική εργασία αναδιαμορφώνονται** σε μια ψηφιακή οικονομία που δίνει προτεραιότητα στην κλίμακα έναντι της βιωσιμότητας.

Το συνέδριο συγκέντρωσε καλλιτέχνες, επαγγελματίες της μουσικής βιομηχανίας, νομικούς, ακαδημαϊκούς, πολιτικούς, διανοητές και τεχνολόγους γύρω από τρία κεντρικά ερωτήματα:

- Πώς διασφαλίζονται **βιώσιμα και δίκαια μοντέλα αμοιβής** σε μια κορεσμένη ψηφιακή αγορά;
- Πώς προστατεύονται τα δικαιώματα των δημιουργών όταν **τα συστήματα ΤΝ εκπαιδεύονται με το έργο τους το οποίο στη συνέχεια ανταγωνίζονται**;
- Ποιος πρέπει να είναι ο ρόλος του νομοθέτη στη διαμόρφωση **ηθικής ΤΝ και δικαιότερων ψηφιακών αγορών**, χωρίς να εμποδίζεται η καινοτομία;

Ο Ο Παύλος Αντωνιάδης (IRCAM/ Πανεπιστήμιο Ιωαννίνων) αμφισβητεί τα βιομηχανικά πρότυπα της μουσικής παραγωγής μέσω ΤΝ. Ξεπερνώντας το «μαύρο κουτί» των μοντέλων Big Data (πχ Suno ή Udio), εξερευνά τους μηχανισμούς της ΑΙ μέσα από το πρίσμα της “σιωπηρής γνώσης”, της ενσώματης, υποκειμενικής εμπειρίας της μουσικής δημιουργίας..



APOLLON
GREEK MUSICIANS'
COLLECTING SOCIETY



ΠΑΡΟΥΣΙΑΣΕΙΣ | ΠΑΥΛΟΣ ΑΝΤΩΝΙΑΔΗΣ

Οι Μηχανισμοί της Τεχνητής Νοημοσύνης

Αθήνα, 23 Οκτωβρίου 2025
Κέντρο Πολιτισμού Ίδρυμα Σταύρος Νιάρχος

Καλημέρα σε όλους

Ευχαριστώ θερμά τους διοργανωτές, και ιδιαίτερα τον Ντίνο Γεωργούντζο, για την ευγενική τους πρόσκληση. Θα ήθελα πολύ να μοιραστώ τον ίδιο φυσικό χώρο μαζί σας, αλλά αυτή τη στιγμή βρίσκομαι στο Πεκίνο της Κίνας για τη 10η έκδοση του συνεδρίου TENOR, με τίτλο "Notational Intelligence" (Σημειογραφική Νοημοσύνη). Γι' αυτό και η μαγνητοσκοπημένη παρέμβασή μου.

Θα ήθελα να συνεισφέρω στο Πάνελ Α, "Οι μηχανισμοί της Τεχνητής Νοημοσύνης" (The mechanics of AI), παρέχοντας μια εναλλακτική λύση στα τρέχοντα βιομηχανικά πρότυπα παραγωγής μέσω Τεχνητής Νοημοσύνης, όπως υλοποιούνται από εργαλεία όπως το Suno ή το Udio. Το όραμά μου συνεπάγεται τη σύλληψη της μουσικής πράξης ("musicking") ως διαδικασία και όχι ως προϊόν, ένα αφήγημα συνέργειας ανθρώπου-μηχανής αντί για αντικατάσταση του ανθρώπου, και μια τεχνική που δεν κεφαλαιοποιεί απαραίτητα τα μεγάλα δεδομένα (big data) και τις αρχιτεκτονικές βαθιάς μάθησης (deep learning).

Αντιθέτως, η εναλλακτική που παρουσιάζω βασίζεται στην άρρητη γνώση (tacit knowledge). Η άρρητη γνώση ορίζεται ως προσωπική, υποκειμενική και εμπειρική, δύσκολη να καταγραφεί στις συνήθεις μορφές δεδομένων· είναι η τεχνογνωσία (know-how) της εμπειρίας της μουσικής δημιουργίας με άλλους. Αυτή η μορφή γνώσης μας βοηθά να εντοπίσουμε τις βαθιές επιστημικές προκαταλήψεις (epistemic biases) που διατρέχουν τα τρέχοντα βιομηχανικά πρότυπα.

Ορίζω αυτές τις προκαταλήψεις ως κάθε μορφή ρητού ή σιωπηρού, θεωρητικού ή τεχνικού περιορισμού στην αναπαράσταση της γνώσης, με παράλογα, ακαθόριστα και ανεξέλεγκτα αποτελέσματα κατά την αλληλεπίδραση του μοντέλου με τον κόσμο. Στο IRCAM, υπογράψαμε πρόσφατα ένα μανιφέστο για την υπεύθυνη τεχνητή νοημοσύνη στην υπηρεσία της εφευρετικής και ποικιλόμορφης μουσικής δημιουργίας, το οποίο εντοπίζει τα κύρια προβλήματα - οικολογικά, κοινωνικά, οικονομικά και καλλιτεχνικά - του τρέχοντος βιομηχανικού προτύπου.

Ο δικός μου ισχυρισμός είναι ότι όλα αυτά τα προβλήματα παραπέμπουν στον ορισμό των επιστημικών προκαταλήψεων που προανέφερα. Ανιχνεύοντας τα φιλοσοφικά και τεχνικά κριτήρια αυτών των προκαταλήψεων, θα επιχειρήσω μια γρήγορη εισαγωγή στα βασικά της μηχανικής μάθησης (machine learning) και της τεχνητής νοημοσύνης στη μουσική, κατόπιν στα βασικά μοντέλα της εμπορικής μουσικής παραγωγής (generation) και, στη συνέχεια, στις σχετικές προκαταλήψεις και τις εναλλακτικές λύσεις.

Γιατί μια μηχανή χρειάζεται να μαθαίνει από τον κόσμο αντί να εκτελεί σκληρά κωδικοποιημένους κανόνες (hardcoded rules), όπως σε αυτό το πρώτο παράδειγμα που βασίζεται στον πατέρα της υπολογιστικής μουσικής, Max Mathews; Η απάντηση είναι ότι μερικές φορές οι κανόνες μπορεί να είναι δύσκολο να εφαρμοστούν λόγω των πολύπλοκων, μεγάλων και θορυβωδών δεδομένων (noisy data), όπως δείχνει το ακόλουθο παράδειγμα του Γιάννη Ξενάκη.

Σε αυτές τις περιπτώσεις, μπορεί να επιλέξουμε να μάθουμε τους κανόνες από τα δεδομένα, καταλήγοντας να προβλέπουμε και να παράγουμε νέα δεδομένα. Αν συγκρίνουμε τους ορισμούς των μηχανικών με τους ορισμούς των μουσικών, όπως φαίνεται σε αυτή τη διαφάνεια, παρατηρούμε ότι οι μουσικοί μπορεί να χρησιμοποιούν τους αλγορίθμους με απροσδόκητους τρόπους, αξιοποιώντας το δημιουργικό τους δυναμικό.

Οι κλασικοί τύποι επιβλεπόμενης (supervised) ή μη επιβλεπόμενης (unsupervised) μηχανικής μάθησης, αλλά και οι τύποι αρχιτεκτονικών βαθιάς μάθησης αιχμής (state-of-the-art), όπως τα RNNs (Αναδρομικά Νευρωνικά Δίκτυα), VAEs (Μεταβλητοί Αυτόματοι Κωδικοποιητές), GANs (Ανταγωνιστικά Δίκτυα Παραγωγής), Transformers ή μοντέλα διάχυσης (diffusion models), δημιουργούν τα πιο σημαντικά συστήματα που κυκλοφορούν. Μοιράζονται βασικά σύνολα δεδομένων (datasets) όπως MIDI, ήχο και φασματογραφήματα (spectrograms), αλλά

συχνά αγνοούν πτυχές της μουσικής πράξης όπως η ενσωμάτωση (embodiment) και η διαδραστικότητα.

Η κεφαλαιοποίησή τους στα μαζικά δεδομένα (big data) συνδέεται όχι μόνο με κοινωνικά προβλήματα, όπως η μικρο-εργασία (micro work) που λαμβάνει χώρα σε καταυλισμούς προσφύγων, όπου οι πρόσφυγες επισημαίνουν (label) δεδομένα για εξευτελιστικές αμοιβές, αλλά και με μια βαθιά προβληματική προσέγγιση του χρόνου και της αλλαγής, καθώς αυτά τα μοντέλα ασχολούνται με αφαιρέσεις του χρόνου παρά με τη δυναμική αλλαγή. Είτε πρόκειται για γραμμικές είτε για μη γραμμικές προσεγγίσεις της χρονικότητας, ο χρόνος ως ένας ανοιχτός ορίζοντας απουσιάζει από αυτά τα μοντέλα. Στο βιβλίο του "Logiciel", ο φιλόσοφος AA Cavia περιγράφει τα βαθιά νευρωνικά δίκτυα ως τοπολογικές δομές με τρεις λειτουργίες: κωδικοποίηση (encoding), ενσωμάτωση (embedding) και αποκωδικοποίηση (decoding).

Η κωδικοποίηση του κόσμου σε δεδομένα είναι κρίσιμη για τους τύπους δεδομένων που συζητήθηκαν παραπάνω. Στην ενσωμάτωση, τα δεδομένα μετασχηματίζονται μέσω κρυφών επιπέδων (hidden layers) που επιτρέπουν την ανίχνευση μοτίβων σε διαφορετικές ιεραρχικές κλίμακες ή επίπεδα αφαίρεσης του λανθάνοντος χώρου (latent space). Η αποκωδικοποίηση της δομής σε ένα αποτέλεσμα (output) είναι μια στοχαστική κατανομή η οποία μπορεί να αντιστοιχιστεί άμεσα στο επίπεδο εισόδου (input layer). Επιστημική βία, λοιπόν, λαμβάνει χώρα και στα τρία στάδια: στο πώς ο κόσμος κωδικοποιείται, ενσωματώνεται και αποκωδικοποιείται στα DNNs (Βαθιά Νευρωνικά Δίκτυα).

Υπάρχει όμως ένα ακόμη βαθύτερο πρόβλημα, αν μου επιτρέπετε το λογοπαίγνιο: το σημασιολογικό τυφλό σημείο της δημιουργίας νοήματος σε αυτά τα μοντέλα, τα οποία λειτουργούν με βάση συσχετίσεις (correlations) και όχι αιτιακές σχέσεις (causations). Η Judea Pearl έχει αναλύσει τα θεωρητικά όρια αυτών των μοντέλων. Μπορούν να συλλογιστούν μόνο σε επίπεδο συσχέτισης, αλλά όχι σε επίπεδο παρέμβασης ή αντιπαραδειγμάτων (counterfactuals). "Τι θα συμβεί αν κάνω το X;" ή "Τι θα είχε συμβεί αν τα δεδομένα εισόδου μου δεν ήταν X αλλά Y;". Αυτό είναι κάτι που μόνο οι άνθρωποι μπορούν να κάνουν: να φαντάζονται διαφορετικούς κόσμους και να παρεμβαίνουν για να τους πραγματοποιήσουν. Τα μοντέλα, τα τρέχοντα μοντέλα, δεν μπορούν να το κάνουν αυτό. Η Marie Shanahan υπογραμμίζει ένα άλλο πρόβλημα, τον λανθάνοντα ανθρωπομορφισμό των LLMs (Μεγάλων

Γλωσσικών Μοντέλων), ο οποίος ενσωματώνεται μέσω της άκριτης χρήσης επιστημικών όρων όπως "σκέφτεται", "γνωρίζει", "πιστεύει".

Στο ίδιο πλαίσιο, τα μοντέλα διάχυσης μουσικής (music diffusion models) δεν συνθέτουν ούτε εκτελούν μουσική, αλλά μάλλον ανασυνδυάζουν -όσο πειστικά κι αν φαίνεται- αυτά τα ψηφία πληροφορίας για να παράγουν στατιστικά ομοιώματα (simulacra). Στο έργο μου από το 2013, συμπεριλαμβανομένης της ανάπτυξης του συστήματος GestCOM, έχω υιοθετήσει με ενθουσιασμό στοχαστικούς σχεδιασμούς, μοντελοποίηση βασισμένη στη μηχανική μάθηση και, πιο πρόσφατα, συστήματα AI όπως το SOMAX II.

Πέρα από τις καθαρά τεχνικές διαφορές τέτοιων εργαλείων από τα βιομηχανικά πρότυπα σε όλα τα στάδια της επεξεργασίας δεδομένων (κωδικοποίηση, ενσωμάτωση και αποκωδικοποίηση), το γεγονός ότι συχνά βασίζονται σε μεμονωμένες χειρονομίες (one-shot gestures) αντί για Μαζικά Δεδομένα MIDI, ήχου και φασματογραφήμάτων (Big Data), ή το γεγονός ότι χρησιμοποιούν τη μηχανική ακρόαση (machine listening) και κύκλους ανάδρασης (feedback circles) της αλληλεπίδρασης ανθρώπου-μηχανής ως προαπαιτούμενο για τη μουσική παραγωγή, και ότι είναι κατά το μεγαλύτερο μέρος εξηγήσιμα (explainable AI) σε αντίθεση με το φαινόμενο του "μαύρου κουτιού" (black box) της βαθιάς μάθησης - τέτοιες προσπάθειες κεφαλαιοποιούν μια πολλαπλότητα αναπαράστασης και αλληλεπίδρασης, επιχειρώντας να συλλάβουν τη φευγαλέα κοινωνική και διαδραστική φύση της μουσικής δημιουργίας που ζει εξίσου στις παρτιτούρες, τις ηχογραφήσεις, τα βίντεο και, πάνω απ' όλα, στη ζωντανή εκτέλεση.

Με τον Mikhail Malt, δημιουργούμε διαλόγους μεταξύ ανθρώπου και μηχανής βασισμένους σε αμοιβαίες διαδικασίες εκπαίδευσης και παραγωγής αποτελεσμάτων, οι οποίες, αν και παράγονται από καλά καθορισμένα σύνολα δεδομένων, ανακτούν τη διαίσθηση ενός πραγματικού αυτοσχεδιασμού, έστω και αν βασίζονται σε κάτι τόσο περίπλοκο όσο η "Ευρυάλη" του Ιάννη Ξενάκη. [Μουσική] Αυτό ήταν μάλλον το δεύτερο παράδειγμά μου, όπου δημιουργώ νέες ευέλικτες αντιστοιχίσεις (mappings) μεταξύ της μουσικής ηχογράφησης και του συστήματος αυτόματου αυτοσχεδιασμού. Και εδώ είμαστε ο Mikhail Malt και εγώ παίζοντας Ξενάκη [Μουσική].

Σε άλλες περιπτώσεις πολυμεσικής σύνθεσης, το "Nimen shì chóngzi" (στα Κινεζικά "Είστε Έντομα"), είναι ένα έργο που στοχάζεται πάνω σε προβλήματα πατρότητας, αυθεντικότητας και εξουσίας στη μουσική Τεχνητή Νοημοσύνη, και

συγκεκριμένα στην ανησυχητική πιθανότητα της ανθρώπινης κατωτερότητας - το γεγονός ότι είμαστε όλοι εν δυνάμει "έντομα" (bugs) με τις πολλαπλές συνυποδηλώσεις του όρου, και ότι ίσως χρειαστεί να αποδεχτούμε μια κατώτερη θέση σε μια νέα ή όχι και τόσο νέα ιεραρχία.

Κλείνοντας, οι τεχνικές λεπτομέρειες της μουσικής ΤΝ δεν είναι απλώς μέσα που μπορούν να χρησιμοποιηθούν για θετικούς ή αρνητικούς σκοπούς, όπως λέει η συνηθισμένη κοινοτοπία. Αντιθέτως, φέρουν βαθύτερα νοήματα σχετικά με τον έλεγχο και τον χρόνο. Κωδικοποιούν, χειραγωγούν και αποκωδικοποιούν όχι μόνο δεδομένα, αλλά τον ίδιο τον κόσμο. Μια διαφορετική μουσική ΤΝ είναι εφικτή, μια ΤΝ που αποκαθιστά τη δυνατότητα παρεμβάσεων και φαντασίωσης πολλαπλών μελλοντικών σεναρίων. Σας ευχαριστώ.

**Παύλος Αντωνιάδης,
IRCAM/Πανεπιστήμιο Ιωαννίνων**

