

## ***AudioScout: Συλλογή ηχητικών δεδομένων στο πεδίο***

Ενές Αχμέτ Κεχαγιά  
Ερευνητής,  
Τμήμα Κινηματογράφου, ΑΠΘ  
enesachm@film.auth.gr

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

Χρήστος Γούσιος  
Αν. Καθηγητής,  
Τμήμα Κινηματογράφου, ΑΠΘ  
cgoussios@film.auth.gr

### **ΠΕΡΙΛΗΨΗ**

Η παρούσα εργασία παρουσιάζει το AudioScout, μια εφαρμογή διαχείρισης ηχητικού υλικού που στοχεύει να αναβαθμίσει την εργασία των sound designers. Η εφαρμογή είναι σχεδιασμένη για να επιτρέπει την αποτελεσματική καταγραφή, οργάνωση, επεξεργασία και αναπαραγωγή ηχητικών δεδομένων. Οι sound designers αντιμετωπίζουν την πρόκληση να αποθηκεύουν και να διαχειρίζονται πολλές ηχογραφήσεις, προκειμένου να δημιουργήσουν το σωστό ηχητικό υπόβαθρο για κάθε παραγωγή. Το AudioScout προσφέρει μια διαισθητική διεπαφή που επιτρέπει την εύκολη ηχογράφηση σε υψηλή ποιότητα όπως επίσης παρέχει τη δυνατότητα ταξινόμησης των αρχείων βάσει μεταδεδομένων όπως τοποθεσία, χρονοσφραγίδα, και είδος ήχου. Στην εφαρμογή θα ενσωματωθούν πιο προηγμένες τεχνολογίες, όπως αυτοματοποιημένη επεξεργασία και αναγνώριση ήχων με βάση τεχνητή νοημοσύνη, καθώς και διασύνδεση με άλλα εργαλεία παραγωγής, αυξάνοντας τη δυναμική και την παραγωγικότητα των επαγγελματιών του ήχου.

### ***AudioScout: Sound Data Collection in the Field***

#### **ABSTRACT**

*This paper introduces AudioScout, a sound management application designed to upgrade the work of sound designers. The application facilitates efficient recording, organization, editing, and playback of audio data. Sound designers face the challenge of storing and managing numerous recordings to create the appropriate audio backdrop for each production. AudioScout provides solutions through an intuitive interface that enables easy, high-quality recording, with capabilities for categorizing files based on metadata such as location, timestamp, and type of sound. The application will integrate more advanced technologies, such as automated processing and sound recognition based on artificial intelligence, as well as interconnectivity with other production tools, enhancing the dynamics and productivity of sound professionals.*

## 1 Εισαγωγή

Στον σύγχρονο κόσμο της κινηματογραφικής παραγωγής αλλά και της παραγωγής οπτικοακουστικού περιεχομένου, γενικότερα, ο ήχος έχει καθιερωθεί ως ένας από τους πυλώνες που επηρεάζουν τη συνολική ποιότητα και την αίσθηση της τελικής παραγωγής. Ο ρόλος του sound designer είναι σημαντικός, καθώς είναι υπεύθυνος για τη δημιουργία, τη σύλληψη και την επεξεργασία των ηχητικών στοιχείων που θα συμβάλλουν στην αφήγηση της ιστορίας. Παρά την καθοριστική συμβολή τους, οι sound designers συχνά αντιμετωπίζουν προκλήσεις που σχετίζονται με τη συλλογή και οργάνωση ηχητικού υλικού, ειδικά όταν απαιτείται να καταγράψουν ήχους στο πεδίο για χρήση σε συγκεκριμένα project στα οποία εργάζονται ή και για μελλοντική χρήση.

Η τεχνολογία έχει προχωρήσει αρκετά, προσφέροντας προηγμένες λύσεις υλικολογισμικού για την καταγραφή και επεξεργασία ήχου [1], ωστόσο, υπάρχει έλλειψη ενός εργαλείου που να επιτρέπει την εύκολη και αποτελεσματική συλλογή, καταγραφή και αναζήτηση ηχοτοπίων. Σε αυτό το πλαίσιο, προτείνεται η ανάπτυξη της εφαρμογής *AudioScout*, μιας προηγμένης, εύχρηστης και πλήρως ενσωματωμένης λύσης για τους επαγγελματίες του ήχου. Το *AudioScout* στοχεύει στην παροχή ενός αξιόπιστου και διαισθητικού εργαλείου που επιτρέπει την άμεση ηχογράφηση, την αποθήκευση με συνοδευτικά μεταδεδομένα (όπως τοποθεσία και χρονοσφραγίδα), και την εύκολη ανάκτηση ηχητικών αρχείων για χρήση σε μελλοντικές παραγωγές.

Η παρούσα εργασία θα παρουσιάσει τη σχεδίαση και τις δυνατότητες της εφαρμογής, εστιάζοντας στην προστιθέμενη αξία που προσφέρει στον τομέα του sound design.

## 2 Πλαίσιο

Το sound design αναφέρεται στην τέχνη και πρακτική της δημιουργίας ηχητικών τοπίων και ηχητικών εφέ για διάφορες εφαρμογές, όπως ταινίες, τηλεοπτικές εκπομπές, θεατρικές παραστάσεις και βιντεοπαιχνίδια. Ο ρόλος του sound designer είναι να επιλέγει, να επεξεργάζεται, και να συνθέτει τον ήχο με τρόπο που να υποστηρίζει και να εμπλουτίζει την οπτική αφήγηση. Η ηχογράφηση στον κινηματογράφο είναι κρίσιμη καθώς τα ηχητικά στοιχεία παίζουν έναν καθοριστικό ρόλο στη δημιουργία της ατμόσφαιρας και την εμβάθυνση της εμπειρίας του θεατή. Ηχογραφήσεις υψηλής ποιότητας μπορούν να μεταφέρουν τον θεατή σε έναν άλλο κόσμο, να βελτιώσουν την αίσθηση της πραγματικότητας [2].

Ένα από τα βασικότερα στοιχεία του συνόλου της ηχητικής εμπειρίας σε κάθε οπτικοακουστική παραγωγή, αποτελούν οι ατμόσφαιρες, γνωστές και ως περιβάλλοντα ή και φόντα (atmospheres, ambiences, backgrounds). Αυτό που ενισχύει τη σημαντικότητα των ήχων αυτών, είναι ακριβώς το γεγονός ότι περνούν - και πρέπει να περνούν- απαρατήρητοι, και οφείλουν να είναι άρρηκτα συνδεδεμένοι με το τοπίο/εικόνα/σύμπαν που παρουσιάζεται, ή αιτιολογημένοι από αυτό. Από την

άλλη βέβαια, είναι ακριβώς αυτοί οι ήχοι που συμβάλλουν και ενισχύουν την πειθώ των εικόνων και εξασφαλίζουν την απερίσπαστη θέαση ή και συμμετοχή. (δηλαδή εξαρτώνται από τις εικόνες, αλλά και οι εικόνες από αυτούς, καθώς τις ολοκληρώνουν). Η εφαρμογή τους και ο τρόπος χρήσης τους ποικίλει, με αποκορύφωμα τη χρήση τους στις κινηματογραφικές παραγωγές, όπου απαιτείται ένα πλήθος ατμοσφαιρών σε κάθε διαφορετική σκηνή, με σκοπό την ηχητικά αφηγηματική πληρότητα [3].

Στην κινηματογραφική αίθουσα, το ηχοσύστημα υποστηρίζει τον πολυκάναλο περιφερειακό ήχο (surround sound και atmos), που παρέχει τη δυνατότητα υποστήριξης της αφήγησης με πολλαπλούς τρόπους, με την τοποθέτηση των διαφορετικών ήχων στο χώρο, άλλοτε δημιουργώντας αληθοφάνεια, άλλοτε ενισχύοντας την εμπύθιση και τη συγκέντρωση στην παρακολούθηση, είτε ενισχύοντας ηχητικά την εξιστόρηση κ.ο.κ. Η τοποθέτηση των ήχων αυτών στο σύνολο του ηχητικού σύμπαντος της εκάστοτε παραγωγής, πραγματοποιείται κατά τη μεταπαραγωγή του ήχου (sound post production), στο στάδιο του ηχητικού σχεδιασμού [4]. Οι εικόνες είναι καθοδηγητικές για τους περισσότερους από τους ήχους αυτούς, αλλά και η φαντασία του κάθε επαγγελματία, συμβάλλει στη σύνθεσή τους. Για παράδειγμα, η εικόνα μιας παραλίας, οδηγεί άμεσα στη χρήση ήχων της θάλασσας, επιπρόσθετα όμως μπορεί να υπάρχει κι ο ήχος ενός μη ορατού αεροπλάνου, που διασχίζει τον ουρανό πάνω από την παραλία αυτή, ή κι ο ήχος ενός επίσης μη ορατού δρόμου, που διέρχεται κοντά από αυτή την παραλία [1]. Οι πηγές των ήχων αυτών είναι ποικίλες και η προέλευσή τους πολλαπλή: υπάρχει πλήθος βιβλιοθηκών ήχου, προς πώληση αλλά και δωρεάν, cloud libraries, αλλά και προσωπικές βιβλιοθήκες του κάθε sound designer. Κάποιες φορές επίσης, μπορεί κάποια παραγωγή να παραγγείλει χαρακτηριστικούς ήχους ατμοσφαιρών από συγκεκριμένο τόπο, ώστε να υπάρχει συνέπεια, κυρίως σε περιπτώσεις όπου ακούγονται και φωνές. Για παράδειγμα, μια παραγωγή που παρουσιάζει ένα εστιατόριο στην Ελλάδα, απαιτεί να ακούγεται και το πλήθος που παρευρίσκεται εκεί το οποίο και μιλάει ελληνικά. Αυτό που είναι σίγουρο είναι ότι όλοι όσοι ασχολούνται με τον ηχητικό σχεδιασμό, ηχογραφούν διαρκώς ήχους τους οποίους και διατηρούν σε προσωπικές τους βιβλιοθήκες.

Η οργάνωση αυτών των ήχων, όπως είναι αναμενόμενο, είναι αρκετά διαφορετική από επαγγελματία σε επαγγελματία, παρά το γεγονός ότι υπάρχουν αρκετά πρωτόκολλα που κυκλοφορούν για τον ορθό τρόπο που πρέπει να ονομάζονται οι ήχοι. Βιβλιοθηκάριοι ήχων προτείνουν τρόπους ονοματοδοσίας, που εξασφαλίζουν την εύκολη προσπέλαση και εύρεση των απαιτούμενων ήχων. Επομένως, η δημιουργία εργαλείων που να εκμεταλλεύονται τις δυνατότητες της τεχνολογίας με στόχο να εξυπηρετούν τους επαγγελματίες προσπαθώντας να συστηματοποιήσουν και να τυποποιήσουν τις πολυποικίλες ροές εργασίας και να ενοποιήσουν τα πρωτόκολλα παρουσιάζει ενδιαφέρον. Σε αυτό το πλαίσιο εντάσσεται και το concept της συγκεκριμένης εργασίας, όπου προσπαθεί να εκμεταλλευτεί σύγχρονες τεχνολογίες όπως τα smartphones, τα φορητά συστήματα ηχοκαταγραφής, την τεχνητή νοημοσύνη και τις υπηρεσίες cloud ώστε να προτείνει μια διαισθητική και εύχρηστη λύση λογισμικού που να αναβαθμίζει την εμπειρία

χρήσης όσον αφορά τη συλλογή, οργάνωση και αναζήτηση ηχητικών καταγραφών, και πιο συγκεκριμένα, περιβαλλόντων/φόντων.

### 3 Ανάπτυξη Εφαρμογής

#### 3.1 Τεχνικές Προδιαγραφές

Η εφαρμογή AudioScout αναπτύσσεται ως Progressive Web App (PWA) για να εξασφαλίσει εύκολη πρόσβαση και ευελιξία στη χρήση, συμβατή τόσο με iOS όσο και με Android. Θα προσφέρει τοπική αποθήκευση δεδομένων, με την επιλογή για συγχρονισμό σε cloud για ασφαλή αποθήκευση και διαχείριση των αντιγράφων ασφαλείας. Τα δεδομένα αποθηκεύονται online, ενώ η διεπαφή χρήστη θα είναι απλή και καθαρή, σχεδιασμένη για να διευκολύνει τη χρήση σε φορητές υπολογιστικές συσκευές.

#### 3.2 Απαιτήσεις

Η εφαρμογή σχεδιάζεται για να παρέχει εύκολη και διαισθητική ηχογράφηση με υψηλή ποιότητα και αποδοτικότητα. Οι βασικές αρχές σχεδίασης συνοψίζονται παρακάτω:

- **Εύκολη Ηχογράφηση:** Η εφαρμογή θα πρέπει να παρέχει έναν απλό και καθαρό τρόπο ηχογράφησης που να επιτρέπει στους χρήστες να ξεκινούν την ηχογράφηση με μόνο ένα κλικ, ακολουθώντας τη λογική της κάμερας ενός κινητού τηλεφώνου.
- **Μεταδεδομένα Ηχογραφήσεων:** Κάθε ηχογράφηση θα πρέπει να συλλέγει αυτόματα μεταδεδομένα όπως τοποθεσία, ημερομηνία, και ώρα της ηχογράφησης, καθώς και η δυνατότητα προσθήκης περιγραφών του περιβάλλοντος ήχου και άλλων σημειώσεων από τον χρήστη.
- **Οργάνωση και Αναζήτηση:** Θα πρέπει να προσφέρονται εκτεταμένες δυνατότητες οργάνωσης των ηχογραφήσεων ανά κατηγορία, τοποθεσία, ή ημερομηνία, με λειτουργία αναζήτησης για τη γρήγορη εύρεση συγκεκριμένων αρχείων.
- **Εξαγωγή και Κοινοποίηση:** Προσφορά δυνατοτήτων για εξαγωγή και κοινοποίηση ηχογραφήσεων σε διάφορες μορφές και μέσω διάφορων πλατφορμών, ώστε οι χρήστες να μπορούν να τις χρησιμοποιούν σε άλλες εφαρμογές ή να τις αποστέλλουν σε συνεργάτες.

#### 3.3 Σχεδιασμός

Ο σχεδιασμός της εφαρμογής βασίζεται σε μια ευέλικτη αρχιτεκτονική που να καλύπτει τις ανάγκες του χρήστη. Ανώτερος στόχος είναι να ενσωματωθούν ολόένα και περισσότερες λειτουργίες, όπως η επεξεργασία και η αποθρομβοποίηση των ηχητικών δεδομένων. Η διεπαφή σχεδιάζεται με στόχο την ελαχιστοποίηση της περιπλοκότητας, επιτρέποντας στους χρήστες να προβαίνουν στην ηχογράφηση και τη διαχείριση του υλικού με λίγα μόνο βήματα. Η ανάπτυξη λαμβάνει υπόψη τις απαιτήσεις για διασφάλιση της ποιότητας του ήχου και της γενικότερης εμπειρίας του χρήστη. Στην επόμενη εικόνα παρουσιάζεται ένα πρώιμο mockup που της κεντρικής οθόνης ηχογράφησης της εφαρμογής.



Εικόνα 1: Προτότυπο χαμηλής πιστότητας της κεντρικής οθόνης ηχογράφησης της εφαρμογής

Οι βασικές αρχές σχεδιασμού για το AudioScout αποτελούν τον θεμέλιο λίθο για την ανάπτυξη μιας αποδοτικής και ευχάριστης εμπειρίας χρήστη. Στοχεύουν στο να δημιουργήσουν ένα σύστημα που είναι εύχρηστο, αυτόματο και προσβάσιμο από παντού, εξυπηρετώντας τον χρήστη με τον καλύτερο δυνατό τρόπο. Οι τρεις βασικές αρχές είναι οι **Intuitive**, **Inputless**, και **Ubiquitous**.

Η αρχή **Intuitive** εξασφαλίζει ότι η εφαρμογή θα είναι εύχρηστη και διαισθητική. Ο απλός σχεδιασμός, η εύκολη πλοήγηση και η γρήγορη απόκριση στις ενέργειες του χρήστη επιτρέπουν μια φυσική αλληλεπίδραση με την εφαρμογή, χωρίς την ανάγκη περίπλοκων οδηγιών ή εκμάθησης. Η αρχή **Inputless** εστιάζει στην ελαχιστοποίηση των απαιτήσεων από τον χρήστη για εισαγωγή δεδομένων. Μέσω της αυτοματοποιημένης επισημείωσης δεδομένων, της μετατροπής ομιλίας σε κείμενο (STT) και της χωροχρονικής επισήμανσης, η εφαρμογή προσφέρει μια πιο έξυπνη και αποδοτική εμπειρία, επιτρέποντας στο σύστημα να εντοπίζει και να καταγράφει πληροφορίες αυτόματα, με ελάχιστη παρέμβαση από τον χρήστη. Τέλος, η αρχή **Ubiquitous** εξασφαλίζει την καθολική διαθεσιμότητα της εφαρμογής σε όλες τις πλατφόρμες. Μέσω της χρήσης τεχνολογιών web, το AudioScout αποτελεί μία progressive web application (PWA), με αποτέλεσμα να είναι συμβατή με περιβάλλοντα iOS, Android και web, και καθώς υποστηρίζει την αποθήκευση στο cloud, επιτρέπει την άμεση πρόσβαση του χρήστη στα δεδομένα του από οπουδήποτε.

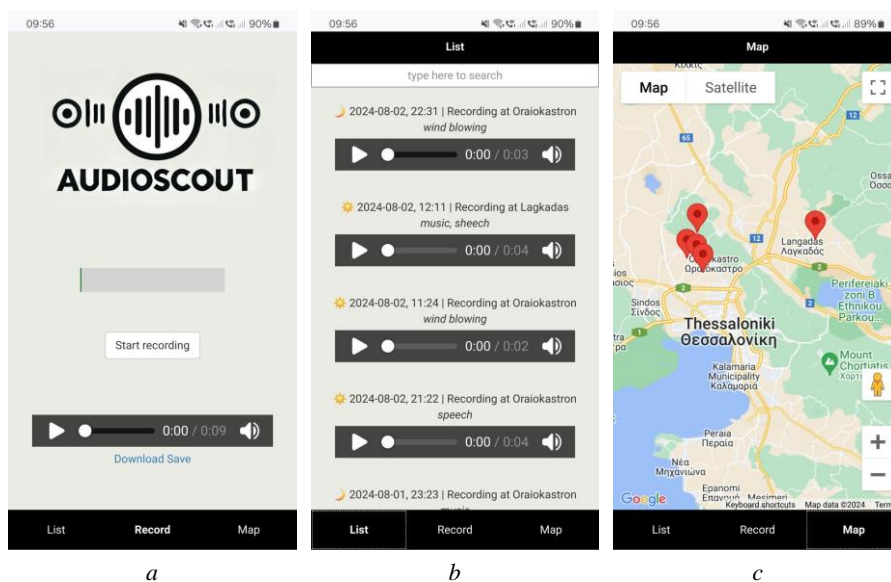
### 3.4 Τεχνολογίες και Εργαλεία Ανάπτυξης

Όπως προαναφέρθηκε, για την ανάπτυξη της εφαρμογής χρησιμοποιούνται σύγχρονες τεχνολογίες web ανάπτυξης, όπως HTML, CSS, και JavaScript. Η λογική επεξεργασίας της εφαρμογής βασίζεται στη JavaScript, αξιοποιώντας τη μεγάλη παρακαταθήκη αλγορίθμων που έχουν υλοποιηθεί στη συγκεκριμένη γλώσσα, αλλά και στη δυνατότητα αξιοποίησης μοντέλων μηχανικής μάθησης (και βαθιάς μάθησης), όπως και τη δυνατότητα επικοινωνίας με εξωτερικές υπηρεσίες web (web

services), στοιχεία που επιταχύνουν την ανάπτυξη. Για τη διαχείριση δεδομένων χρησιμοποιείται μια σχεσιακή βάση δεδομένων MySQL, διευκολύνοντας τη διαχείριση δεδομένων. Τέλος, όλα τα δεδομένα του χρήστη αποθηκεύονται στο cloud, ενώ η κρυπτογράφησή τους είναι υπό διερεύνηση, ώστε να αναβαθμιστεί η πυχύ της ασφάλειας.

### 3.5 Λειτουργικότητα και Χαρακτηριστικά

Η εφαρμογή AudioScout έχει σχεδιαστεί για να προσφέρει μια σειρά από λειτουργίες που καθιστούν την ηχογράφιση, την αποθήκευση, την επεξεργασία και την ανάκτηση ηχητικών αρχείων αποτελεσματική και ιδιαίτερα εύχρηστη για sound designers.



Εικόνα 2: Screenshots από τις βασικές οθόνες της εφαρμογής AudioScout: a) Επιλογή ποιότητας και μορφής ήχου για άμεσες ηχογραφήσεις, b) κατηγοριοποίηση και ετικετοποίηση αρχείων για εύκολη αναζήτηση, c) γεωγραφική αναζήτηση και προβολή ηχογραφήσεων στον χάρτη.

#### i. Άμεση Ηχογράφιση

Οι χρήστες μπορούν να ξεκινήσουν ηχογραφήσεις αμέσως μόλις ανοίξουν την εφαρμογή, χωρίς να χρειάζεται να περιηγηθούν μέσα από περίπλοκες ρυθμίσεις ή επιλογές. Η εφαρμογή είναι σχεδιασμένη για να καταγράφει ηχοτοπία με ευκολία, χρησιμοποιώντας προεπιλεγμένες ρυθμίσεις ήχου που εξασφαλίζουν μικρό μέγεθος αρχείων, αλλά και επαρκή ποιότητα του ηχητικού υλικού.

#### ii. Οργάνωση και Ταξινόμηση Ηχητικών Αρχείων

Η εφαρμογή προσφέρει δυνατότητες οργάνωσης, επιτρέποντας στους χρήστες να καταχωρούν ηχητικά αρχεία σε κατηγορίες, να προσθέτουν ετικέτες και να

διαχειρίζονται τα μεταδεδομένα όπως τοποθεσία και χρονοσφραγίδα. Αυτό καθιστά την αναζήτηση και την ανάκτηση ηχητικών αρχείων γρήγορη και αποτελεσματική.

### iii. Απεικόνιση Ηχογραφήσεων σε Χάρτη

Η εφαρμογή ενσωματώνει μια δυναμική λειτουργία απεικόνισης ηχογραφήσεων σε χάρτη, όπου κάθε ηχογράφιση συνοδεύεται από μια πινέζα στην ακριβή τοποθεσία που έγινε. Οι χρήστες μπορούν να πλοηγηθούν εύκολα στον χάρτη για να εξερευνήσουν ηχογραφήσεις σε διάφορες τοποθεσίες, προσφέροντας μια οπτική και διαδραστική εμπειρία αναζήτησης. Αυτή η λειτουργία επιτρέπει επίσης την αναζήτηση ηχογραφήσεων μέσω της γεωγραφικής τους θέσης, καθιστώντας την επιλογή και ανάκτηση των αρχείων πιο άμεση και εύκολη.

## 4 Μελλοντικές εξελίξεις

Η ανάγκη για συνεχή βελτίωση και επέκταση των λειτουργιών είναι πάντα δεδομένη, ώστε μια εφαρμογή λογισμικού παραμείνει στην κορυφή της τεχνολογίας και να ικανοποιεί τις ανάγκες των επαγγελματιών του ήχου. Όσον αφορά τη συγκεκριμένη περίπτωση, αρχικά, στο AudioScout μπορούν να ενσωματωθούν εφέ για την επεξεργασία του ήχου, όπως ισοστάθμιση, φιλτράρισμα θορύβου, και επεξεργασία δυναμικών για τη βελτιστοποίηση του ηχητικού σήματος. Έπειτα, χρησιμοποιώντας τεχνικές τεχνητής νοημοσύνης, η εφαρμογή θα μπορεί να αναγνωρίσει και να καταγράψει ηχητικά περιβάλλοντα, προσφέροντας στους χρήστες μια πλούσια βιβλιοθήκη από επισημειωμένα ηχητικά στιγμιότυπα η οποία μπορεί να χρησιμοποιηθεί για αναφορές ή άμεση χρήση σε παραγωγές. Τέλος, η εφαρμογή θα μπορεί να υποστηρίξει εκτεταμένες δυνατότητες συνεργασίας, επιτρέποντας σε πολλαπλούς χρήστες να συνεργάζονται σε ηχητικά πρότζεκτ, να μοιράζονται αρχεία και να συντονίζουν τις εργασίες τους εντός της πλατφόρμας, αυξάνοντας την παραγωγικότητα και τη δημιουργική συνεργασία.

## 5 Αναφορές

2. Vrysis, L., Tsipas, N., Dimoulas, C., & Papanikolaou, G. (2015). Mobile audio intelligence: From real-time segmentation to crowdsourced semantics. In *Audio Mostly 2015*.
3. Gousios C. (2018), *Sound & Music in Cinema* [PowerPoint slides], School of Film, AUTh.
4. Schafer, R. M. (1993). *The soundscape: Our sonic environment and the tuning of the world*. Simon and Schuster.
5. Chion, M. (2010). *Ο ήχος στον κινηματογράφο*. Εκδόσεις Πατάκη, Paris France.