



## Σχολή Εφαρμοσμένων Τεχνών

### Ακουστικός Σχεδιασμός και Ψηφιακός Ήχος

#### Διπλωματική Εργασία

«Ανάπτυξη εφαρμογής Android για την αυτόματη αξιολόγηση της ποιότητας ηχογραφήσεων»

Ιωάννης Λελεδάκης

Επιβλέπων καθηγητής: Βρύσης Λάζαρος

Ηράκλειο, Ιούνιος 2024

© Ελληνικό Ανοικτό Πανεπιστήμιο, 2017

Η παρούσα Εργασία καθώς και τα αποτελέσματα αυτής, αποτελούν συνιδιοκτησία του ΕΑΠ και του φοιτητή, ο καθένας από τους οποίους έχει το δικαίωμα ανεξάρτητης χρήσης, αναπαραγωγής και αναδιανομής τους (στο σύνολο ή τμηματικά) για διδακτικούς και ερευνητικούς σκοπούς, σε κάθε περίπτωση αναφέροντας τον τίτλο και το συγγραφέα της Εργασίας καθώς και το όνομα του ΕΑΠ όπου εκπονήθηκε.

«Ανάπτυξη εφαρμογής Android για την αυτόματη αξιολόγηση της ποιότητας ηχογραφήσεων»

Λελεδάκης Ιωάννης

Επιτροπή Επίβλεψης Διπλωματικής Εργασίας

Επιβλέπων Καθηγητής:

Βρύσης Λάζαρος

Αριστοτέλειο Πανεπιστήμιο

Θεσσαλονίκης (ΑΠΘ)

Συν-Επιβλέπων Καθηγητής:

Δημουλάς Χαράλαμπος

Αριστοτέλειο Πανεπιστήμιο

Θεσσαλονίκης (ΑΠΘ)

Ηράκλειο, Ιούνιος 2024

## Περίληψη

Στην εργασία αυτή έγινε η ανάπτυξη μιας εφαρμογής για λειτουργικό σύστημα Android η οποία πραγματοποιεί ηχογράφηση, και από το αρχείο ήχου που προκύπτει, υπολογίζει φασματικά χαρακτηριστικά του ηχογραφημένου σήματος. Από τα χαρακτηριστικά αυτά εξάγονται συμπεράσματα για την ποιότητα της ηχογράφησης με σκοπό την ανίχνευση προβλημάτων, όπως την παρουσία θορύβου, ή και την ύπαρξη παραμόρφωσης ψαλιδισμού. Τα αποτελέσματα παρουσιάζονται στο χρήστη μέσω μίας απλής αλλά εργονομικής διεπαφής χρήστη. Τέλος, πραγματοποιήθηκε μικρής κλίμακας μελέτη ευχρηστίας, όπου μικρός αριθμός χρηστών συμπλήρωσε ερωτηματολόγιο που παρείχαμε.

## Λέξεις – Κλειδιά

Ψηφιακή Επεξεργασία Σήματος, Εξαγωγή Χαρακτηριστικών Ήχου

## **Abstract**

In this work, an application for the Android operating system was developed which performs audio recording, and from the resulting audio file we extract audio features of the recorded signal. From said features we draw conclusions about the quality of the recording, such as the presence and extent of noise in our signal and the existence of clipping distortion. The results are presented to the user through a simple, yet ergonomic user interface. Finally, a small-scale usability study was conducted, where a small number of users completed a questionnaire that we provided.

## **Keywords**

Digital Audio Signal Processing, Audio Feature Extraction, Android Audio

## Περιεχόμενα

Περίληψη .....	4
Abstract .....	5
Περιεχόμενα .....	6
Κατάλογος Εικόνων / Σχημάτων .....	9
1 Εισαγωγή.....	1
1.1 Στόχος της εργασίας.....	2
1.2 Εφαρμογές ψηφιακού ήχου σε συσκευές Android .....	2
1.2.1 Εφαρμογές τύπου εργαλείου.....	2
1.2.2 Εφαρμογές τύπου αναπαραγωγής μουσικής.....	3
1.2.3 Εφαρμογές τύπου Recorder/Editor/DAW .....	4
1.3 Παράμετροι ποιοτικής αξιολόγησης ηχογραφήσεων .....	5
1.3.1 Στάθμη σήματος εισόδου – Ψαλιδισμός κυματομορφής .....	5
1.3.2 Χαρακτηριστικά ακουστικής του χώρου ηχογράφησης .....	6
1.4 Χαρακτηριστικά και μεγέθη που θα χρησιμοποιηθούν .....	6
1.4.1 Ανίχνευση ψαλιδισμού κυματομορφής .....	6
1.4.2 Ανίχνευση παρουσίας θορύβου .....	7
2 Επισκόπηση του πεδίου .....	8
2.1 Εφαρμογές/Εργαλεία αξιολόγησης ηχογραφήσεων – μη αυτόματα .....	8
2.2 Εργαλεία/Ερευνες για την Αυτοματοποιημένη βελτίωση ηχογραφήσεων .....	11
3 Μεθοδολογία και υλοποίηση .....	14
3.1 Το λειτουργικό σύστημα φορητών συσκευών Android.....	14
3.2 Η βιβλιοθήκη ψηφιακής επεξεργασίας σήματος JTransforms .....	15

3.3	Καταγραφή PCM ηχητικών δεδομένων από το μικρόφωνο της κινητής συσκευής	16
3.4	Αναπαραγωγή των αρχείων ηχογραφήσεων.....	19
3.5	Επεξεργασία των ηχογραφημένων αρχείων .....	21
3.6	Πλήρες Διάγραμμα Κλάσεων .....	23
3.7	Μέθοδοι υπολογισμού παραμέτρων αξιολόγησης.....	24
3.7.1	Μέθοδος βαθμολόγησης ηχογράφησης .....	24
3.7.2	Ψαλιδισμός Κυματομορφής (Clipping) .....	26
3.7.3	Φασματική επιπεδότητα.....	27
3.7.4	Συχνοτικό Εύρος (Bandwidth).....	28
3.7.5	Φασματικό Κέντρο Μάζας (Spectral centroid).....	28
3.7.6	Spectral spread .....	29
3.8	Δημιουργία της Εφαρμογής .....	30
3.8.1	Android Studio.....	30
3.8.2	Σχεδίαση διεπαφής χρήστη με τη χρήση των Views .....	31
3.8.3	Γραφική διεπαφή χρήστη .....	33
4	Δοκιμές ευχρηστίας της εφαρμογής .....	36
4.1	Ευχρηστία .....	36
4.2	Αρχές Ευχρηστίας σε Εφαρμογές Android.....	36
4.3	Εργαλεία και Πρακτικές για Βελτίωση της Ευχρηστίας .....	37
4.4	Σχεδιασμός δοκιμών ευχρηστίας για εφαρμογές Android.....	38
4.5	Μέθοδος Likert .....	39
4.6	Ερωτηματολόγιο αξιολόγησης.....	40
4.7	Αποτελέσματα.....	41

4.8	Συζήτηση.....	49
5	Συμπεράσματα και μελλοντικές κατευθύνσεις.....	52
	Βιβλιογραφία .....	54
	Παράρτημα Α: Σύνδεσμος με την πλήρη λίστα κώδικα της εφαρμογής.....	59
	Παράρτημα Β: Απαντήσεις εξειδικευμένων χρηστών.....	60
	Χρήστης Α .....	60
	Χρήστης Β .....	60
	Χρήστης Γ.....	61
	Παράρτημα Γ: Σύνδεσμος με πίνακα αποτελεσμάτων ερωτηματολογίου.....	62



## Κατάλογος Εικόνων / Σχημάτων

Εικόνα 1: Λογότυπο της εφαρμογής Foobar2000 (αριστερά) και της VLC Media Player (δεξιά) .....	3
Εικόνα 2: Λογότυπο της εφαρμογής Spotify (αριστερά) και της Tidal (δεξιά).....	3
Εικόνα 3: Λογότυπο της εφαρμογής DiscDj 3D Music Player (αριστερά) και Virtual DJ Mixer (δεξιά).....	4
Εικόνα 4: Λογότυπο της εφαρμογής nTrack Studio(αριστερά) και wavepad (δεξιά) .....	4
Εικόνα 5: Παράδειγμα ψαλιδισμένης κυματομορφής(‘WHAT IS CLIPPING?   MACKIE’, n.d.) .....	5
Εικόνα 6 .....	8
Εικόνα 7:Διεπαφή της εφαρμογής WaveEditor for Android.....	8
Εικόνα 8 .....	9
Εικόνα 9: Διεπαφή της εφαρμογής FL Studiio Mobile .....	9
Εικόνα 10 .....	10
Εικόνα 11: Διεπαφή της εφαρμογής Audio Evolution Mobile Studio .....	10
Εικόνα 12 .....	11
Εικόνα 13: Τέσσερα στάδια της λειτουργίας της εφαρμογής.....	12
Εικόνα 14 .....	13
Εικόνα 15: Διάγραμμα κλάσης του πακέτου record.....	16
Εικόνα 16: Διάγραμμα κλάσης του πακέτου playback.....	19
Εικόνα 17: Διάγραμμα κλάσης του πακέτου process .....	21
Εικόνα 18: Πλήρες διάγραμμα κλάσεων .....	23

Εικόνα 19: (αριστερά) Γραφική διεπαφή χρήστη της εφαρμογής μας μόλις αυτή έχει ανοίξει .....	34
Εικόνα 20: Εικονίδιο Εφαρμογής Audio Recording Evaluator .....	35
Εικόνα 21: QR code με τον σύνδεσμο της εφαρμογής στο ηλεκτρονικό κατάστημα της GooglePlay της Google.....	35
Εικόνα 22: QR-code με σύνδεσμο για το ερωτηματολόγιο της εργασίας. ....	40
Εικόνα 23: Αποτελέσματα απαντήσεων ερωτηματολογίου .....	41
Εικόνα 24: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την πρώτη ερώτηση του ερωτηματολογίου .....	42
Εικόνα 25:διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την δεύτερη ερώτηση του ερωτηματολογίου .....	42
Εικόνα 26: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την τρίτη ερώτηση του ερωτηματολογίου .....	43
Εικόνα 27: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την τέταρτη ερώτηση του ερωτηματολογίου .....	43
Εικόνα 28: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την πέμπτη ερώτηση του ερωτηματολογίου .....	44
Εικόνα 29: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την έκτη ερώτηση του ερωτηματολογίου .....	44
Εικόνα 30:διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την έβδομη ερώτηση του ερωτηματολογίου .....	45
Εικόνα 31: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την όγδοη ερώτηση του ερωτηματολογίου .....	45
Εικόνα 32: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την ένατη ερώτηση του ερωτηματολογίου .....	46

Εικόνα 33: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την δέκατη ερώτηση του ερωτηματολογίου .....	46
Εικόνα 34: Διάγραμμα πίτας με την κατανομή των συμμετεχόντων στις τρεις ηλικιακές ομάδες που χρησιμοποιήσαμε.....	47
Εικόνα 35: Διάγραμμα πίτας με την κατανομή των συμμετεχόντων με βάση τη φύλο τους. .	47
Εικόνα 36: Διάγραμμα πίτας με την κατανομή των συμμετεχόντων βάσει του μορφωτικού τους επιπέδου. ....	48
Εικόνα 37: Διάγραμμα, με την κατανομή των συμμετεχόντων βάσει της σχετικής τους άνεσης με την τεχνολογία γενικά.....	48
Εικόνα 38: Σύνδεσμος με πίνακα αποτελεσμάτων ερωτηματολογίου .....	62

## 1 Εισαγωγή

Η δυνατότητα αναπαραγωγής, εγγραφής, και επεξεργασίας ήχου από φορητές συσκευές είναι αντικείμενο αναζήτησης και έρευνας από πολύ παλιά, με πληθώρα προϊόντων από πολλές εταιρίες να έχουν εμφανιστεί ανά τα χρόνια προσπαθώντας να καλύψουν την ανάγκη αυτή. Μερικά παραδείγματα αυτών είναι το Walkman που δεχόταν κασέτες, τα διάφορα φορητά CD players, αργότερα με την έλευση του mp3, τα φορητά mp3 players τα οποία βελτίωσαν δραματικά τη φορητότητα συσκευών αναπαραγωγής ήχου με το χαρακτηριστικά μικρό τους μέγεθος σε σχέση με τις παλιότερες συσκευές και την απουσία της προβληματικής αναπαραγωγής υπό την παρουσία δονήσεων. Τα mp3 players έδωσαν με τη σειρά τους τη σκυτάλη στα πρώτα κινητά με δυνατότητα αναπαραγωγής πολυμέσων, και από εκεί και έπειτα με την έλευση των smartphones, όπως το iPhone και τα κινητά Android, και την ενσωμάτωση της πληθώρας των δυνατοτήτων που συνδύαζαν εντός μίας συσκευής οδήγησαν στην εγκατάλειψη των φορητών players.

Στο κομμάτι της ηχογράφησης από φορητές συσκευές η αγορά ήταν περισσότερο προσανατολισμένη στους επαγγελματίες όπως δημοσιογράφοι, που χρησιμοποιούσαν φορητά καταγραφικά κασέτας, τα οποία σε διάφορες παραλλαγές, διατάξεις και μεγέθη χρησιμοποιούνταν και από καλλιτέχνες για πρόχειρη ηχογράφηση μουσικής στην περίπτωση μουσικών, ήχων φύσης από τοπία κλπ. για την εξαγωγή δειγμάτων στην περίπτωση κινηματογραφιστών (συνήθως). Τα καταγραφικά ακολουθώντας και αυτά την εξέλιξη των μέσων αποθήκευσης, από μέσα μαγνητοταινίας σε ψηφιακά μέσα όπως κάρτες μνήμης, σκληρούς δίσκους κλπ., ανάλογα την περίπτωση και την υλοποίηση στην πορεία των χρόνων αύξησαν δραματικά τις δυνατότητες καταγραφής τόσο σε όγκο δεδομένων αλλά και σε φορητότητα συσκευών. Με την ανάπτυξη των Social Media και Social Networks οι ανάγκες του απλού καταναλωτή για ποιοτική καταγραφή εικόνας και ήχου για την παραγωγή πολυμεσικού περιεχομένου για την εκάστοτε πλατφόρμα (YouTube, Facebook, κλπ.) προσέγγισαν σημαντικά αυτές των επαγγελματιών, κατά συνέπεια οι κατασκευαστές ενσωμάτωσαν στις συσκευές smartphone δυνατότητες, τέτοιες ώστε οι ανάγκες σε ποιότητα και αποθήκευση, με τη χρήση αυτοματισμών, ολοένα βελτιούμενων ανάλογα με την

υπολογιστική ισχύ και μνήμη των συσκευών, να καλύπτονται (The Saturday Evening Post, 2023; GTE Media, 2023).

Σε αυτό το κεφάλαιο παρουσιάζουμε τους διάφορους τύπους εφαρμογών android με αντικείμενο τον ήχο και κάποια από τα φαινόμενα και μεγέθη που θα μας απασχολήσουν στη συνέχεια.

## 1.1 Στόχος της εργασίας

Στην εργασία αυτή θα ασχοληθούμε με το κομμάτι της ηχογράφησης σε φορητές συσκευές Android και την αυτόματη αξιολόγηση των ηχογραφήσεων αυτών. Στις δύο παρακάτω ενότητες, θα αναλύσουμε τις παραμέτρους/χαρακτηριστικά του καταγεγραμμένου υλικού με βάση τις οποίες θα γίνει η αξιολόγηση αυτή. Αυτές οι παράμετροι περιλαμβάνουν το φαινόμενο του ψαλιδισμού της κυματομορφής και την ύπαρξη υψηλής στάθμης αντηχήσεων ή περιβαλλοντικού θορύβου. Τέλος, θα αναπτυχθεί ένα αλγοριθμικό πλαίσιο για την αξιολόγηση των ηχογραφήσεων, το οποίο θα λαμβάνει υπόψη αυτές τις παραμέτρους, και θα υλοποιηθεί σε μια εφαρμογή για φορητές συσκευές με λειτουργικό σύστημα Android.

## 1.2 Εφαρμογές ψηφιακού ήχου σε συσκευές Android

Οι εφαρμογές Android για ανάλυση και επεξεργασία, που υπάρχουν αυτή τη στιγμή διαθέσιμες, υπάγονται σε τρεις βασικές κατηγορίες, σε εφαρμογές τύπου «εργαλείο μετρήσεων», εφαρμογές τύπου αναπαραγωγής ήχου, και εφαρμογές τύπου audio editor/sequencer ή DAW. Παρακάτω αναλύουμε κάποια παραδείγματα αυτών.

### 1.2.1 Εφαρμογές τύπου εργαλείου

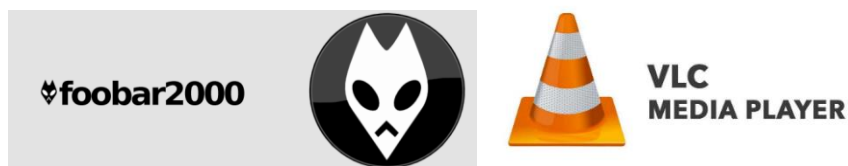
Δύο από τα πιο κοινά παραδείγματα εφαρμογών τύπου εργαλείου είναι οι υλοποιήσεις ηχόμετρων και αναλυτών φάσματος ενώ περιλαμβάνει και εργαλεία όπως το shazam που κάνουν audio fingerprinting για την αναζήτηση τραγουδιών. Τέλος έχουμε και τις εφαρμογές

δοκιμής ηχείων οι οποίες συνδέονται σε ένα ηχοσύστημα, το τροφοδοτούν με σήματα μετρήσεων και με τη χρήση του μικροφώνου της συσκευής πραγματοποιούν διάφορες μετρήσεις για την εξαγωγή συμπερασμάτων

### 1.2.2 Εφαρμογές τύπου αναπαραγωγής μουσικής

Η κατηγορία εφαρμογών αναπαραγωγής μουσικής περιλαμβάνει παραδείγματα όπως:

Το foobar2000 και τον VLC Media Player (για Android, πιο γενικό player και για βίντεο), τα οποία κάνουν αναπαραγωγή κυρίως τοπικών αρχείων με δυνατότητα αναπαραγωγής ροών από δίκτυο ή το διαδίκτυο.



Εικόνα 1: Λογότυπο της εφαρμογής Foobar2000 (αριστερά) και της VLC Media Player (δεξιά)

Αντίστοιχα υπάρχουν εφαρμογές/υπηρεσίες όπου ο χρήστης μπορεί να αναπαράγει αρχεία ήχου μέσω δικτυακών ροών on-demand. Τέτοιες υπηρεσίες περιλαμβάνουν τα:

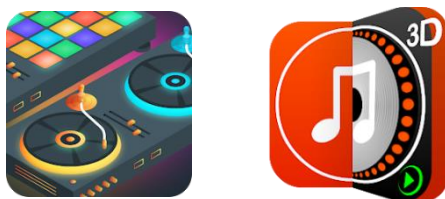
**Spotify**, μια ψηφιακή υπηρεσία μουσικής, podcast και βίντεο που σου δίνει πρόσβαση σε εκατομμύρια τραγούδια και άλλο περιεχόμενο από δημιουργούς από όλο τον κόσμο.



Εικόνα 2: Λογότυπο της εφαρμογής Spotify (αριστερά) και της Tidal (δεξιά)

Το **Tidal**, επί της ουσίας ίδιο αλλά προωθεί ότι παρέχει ροές υψηλής ποιότητας (CD-44100Hz 16bit, υψηλής ανάλυσης FLAC, και Dolby Atmos) στην premium έκδοση.

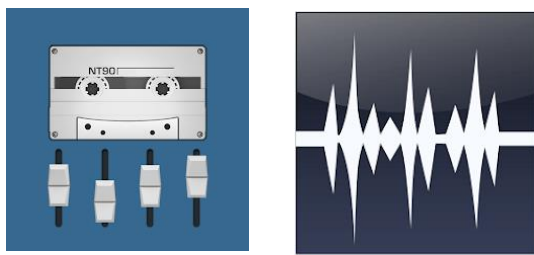
Τέλος εφαρμογές τύπου DJ όπως οι DiscDj 3D Music Player και το Virtual DJ Mixer οι οποίες επαυξάνουν τον πρώτο τύπο με λειτουργίες εφαρμογής εφέ, μίξης, επανάληψης αποσπασμάτων, και εξαγωγής δειγμάτων μεταξύ άλλων ή ακόμα εξομοιώνοντας πλήρως έναν μίκτη DJ.



Εικόνα 3: Λογότυπο της εφαρμογής DiscDj 3D Music Player (αριστερά) και Virtual DJ Mixer (δεξιά)

### 1.2.3 Εφαρμογές τύπου Recorder/Editor/DAW

Τέλος, η Τρίτη κατηγορία αφορά εφαρμογές όπως το nTrack που είναι ουσιαστικά ένα sequencer/DAW, το wavepad που είναι ένας audio editor για Android και iOS.



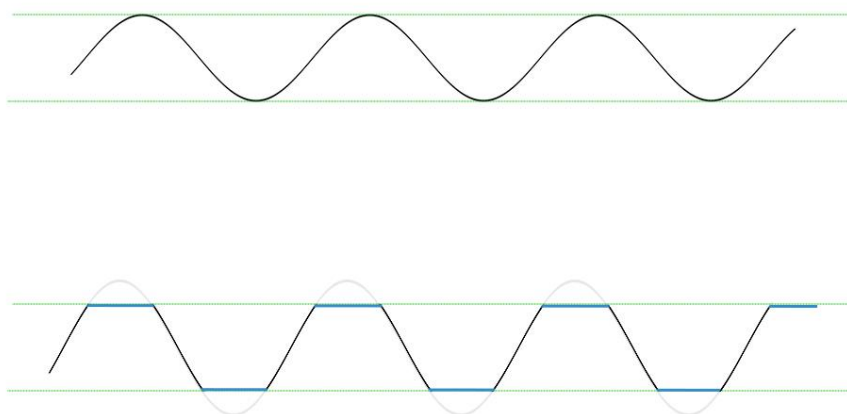
Εικόνα 4: Λογότυπο της εφαρμογής nTrack Studio(αριστερά) και wavepad (δεξιά)

Οι εφαρμογές αυτές έχουν όλες τις δυνατότητες και λειτουργίες ενός DAW με μόνους περιορισμούς/διαφορές, τα χαρακτηριστικά αλληλεπίδρασης της συσκευής με τον χρήστη και τις δυνατότητες συνδεσιμότητας με άλλες συσκευές όπως διεπαφές ήχου.

## 1.3 Παράμετροι ποιοτικής αξιολόγησης ηχογραφήσεων

### 1.3.1 Στάθμη σήματος εισόδου – Ψαλιδισμός κυματομορφής

Οι παράμετροι αυτές περιλαμβάνουν την ύπαρξη ή μη ψαλιδισμού στην είσοδο, κάτι που μπορεί να συμβεί με δύο τρόπους, είτε η κάψα του μικροφώνου λαμβάνει πολύ ισχυρά ηχητικά κύματα από την πηγή που καταγράφουμε, είτε το κέρδος εισόδου της προενίσχυσης της συσκευής που τροφοδοτεί με σήμα το μικρόφωνο είναι πάρα πολύ υψηλό, με αποτέλεσμα να ξεπερνά το δυναμικό εύρος της συσκευής καταγραφής και να «ψαλιδίζεται» η κυματομορφή του σήματος εισόδου. Δηλαδή το σήμα μας είτε καταφτάνει στην προενίσχυση «ψαλιδισμένο» λόγω μιας πολύ δυνατής πηγής, είτε «ψαλιδίζεται» από την προενίσχυση με ακατάλληλη ρύθμιση κέρδους εισόδου (input gain). Οπτικά αυτό φαίνεται σε μία κυματομορφή σαν τετραγωνισμό των κορυφών του σήματος.



**Εικόνα 5: Παράδειγμα ψαλιδισμένης κυματομορφής**(‘WHAT IS CLIPPING? | MACKIE’, n.d.)

Τέλος, να σημειώσουμε ότι ένας συνήθης παράγοντας υποβάθμισης της ποιότητας μίας ηχογράφησης είναι ο θόρυβος από τα ηλεκτρονικά/ηλεκτρολογικά της όποιας συνδεσμολογίας/εγκατάστασης χρησιμοποιείται και οφείλεται συνήθως σε κακές συνδέσεις από χαλασμένα καλώδια ή από βρόχους γείωσης. Στα πλαίσια αυτής της ΔΕ, δεν θα μας απασχολήσουν τέτοιοι θόρυβοι αφού ασχολούμαστε με συσκευές που έχουν ενσωματωμένο μικρόφωνο.



### 1.3.2 Χαρακτηριστικά ακουστικής του χώρου ηχογράφησης

Τα χαρακτηριστικά αυτά περιλαμβάνουν τον θόρυβο βάθους ή περιβαλλοντικό θόρυβο, όπου αν πρόκειται για ανοιχτό χώρο, αυτό αφορά θορύβους όπως οχλαγωγία, άνεμο, και θόρυβο αυτοκινήτων από κοντινό δρόμο μεταξύ άλλων. Αν πρόκειται για κλειστό χώρο, τα ακουστικά χαρακτηριστικά αυτού, όπως ο χρόνος αντήχησης και η στάθμη των αντηχήσεων, παίζουν πολύ σημαντικό ρόλο στην ποιότητα της όποιας ηχογράφησης πραγματοποιηθεί εντός αυτού. Αν, για παράδειγμα, η λήψη δεν γίνεται από επαρκώς κοντινή απόσταση στην πηγή, οι αντηχήσεις ενός, ενδεχομένως πλούσιου σε ανακλαστικές επιφάνειες, δωματίου, να υποβαθμίσουν αρκετά την καταληπτότητα μιας ηχογράφησης ομιλίας.. Άλλα ακουστικά χαρακτηριστικά όπως ρυθμοί του δωματίου επηρεάζουν και αυτοί την και την ποιότητα της ηχογράφησης, συνεισφέροντας στην αλλοίωση της χροιάς της όποιας πηγής ηχογραφείται.

## 1.4 Χαρακτηριστικά και μεγέθη που θα χρησιμοποιηθούν

Για τον έλεγχο της ποιότητας της ηχογράφησης από την οποία προέκυψε το ηχητικό υλικό του αρχείου που επεξεργαζόμαστε, χρησιμοποιούνται η ύπαρξη ψαλιδισμού στο σήμα, το μέτρο της φασματικής επιπεδότητας, το συχνотικό εύρος, το φασματικό κεντροειδές, και η φασματική διασπορά .

### 1.4.1 Ανίχνευση ψαλιδισμού κυματομορφής

Ένα από τα βασικότερα και πιο χαρακτηριστικά σφάλματα κατά την ηχογράφηση μιας οποιασδήποτε πηγής είναι η στάθμη καταγραφής να υπερβαίνει τη μέγιστη στάθμη που μπορεί να καταγράψει το σύστημά μας. Αυτό καταγράφεται στις τιμές των δειγμάτων του ηχητικού υλικού μας με αυτά να παίρνουν τη μέγιστη τιμή κβαντισμού συνεχόμενα για κάποιο χρονικό διάστημα κατά το οποίο τη σήμα εισόδου έφτασε στη βαθμίδα καταγραφής με στάθμη μεγαλύτερη από αυτή που μπορεί να καταγραφεί. Στην περίπτωση αυτή, ελέγχοντας τις τιμές των δειγμάτων για υπέρβαση της μέγιστης τιμής κβαντισμού μπορούμε να καθορίσουμε αν υπάρχει, σε ποιο χρονικό σημείο, και σε τι βαθμό ψαλιδισμός στο σήμα

μας. Παρακάτω στο κεφάλαιο υλοποίηση εμβαθύνουμε περισσότερο στο πρόβλημα αυτό και εξηγούμε πως κινηθήκαμε για την ανίχνευσή του στα πλαίσια της εφαρμογής μας

#### 1.4.2 Ανίχνευση παρουσίας θορύβου

Για την ανίχνευση της παρουσίας θορύβου στο σήμα μας χρησιμοποιούμε το μέτρο της φασματικής επιπεδότητας (Spectral Flatness Measure - SFM), το εύρος ζώνης (Bandwidth), το φασματικό κέντρο μάζας, και τη φασματική διασπορά.

Το SFM παίρνει τιμές από μηδέν έως ένα και όσο πιο κοντά στη μονάδα είναι οι τιμές, τόσο περισσότερο το σήμα μας προσεγγίζει το λευκό θόρυβο. Το εύρος ζώνης, μας πληροφορεί για το πόσο ευρύ είναι το φάσμα του σήματός μας από μία στάθμη και πάνω. Το Spectral Centroid μας πληροφορεί για το που είναι συγκεντρωμένη η περισσότερη ενέργεια του φάσματος του σήματός μας – προς τις υψηλές ή τις χαμηλές συχνότητες δηλαδή, και η φασματική διασπορά συνδεδεμένη με το spectral centroid, μας πληροφορεί για το κατά πόσο απλώνεται η ενέργεια αυτή, μακριά από το τονικό κέντρο του φάσματος.

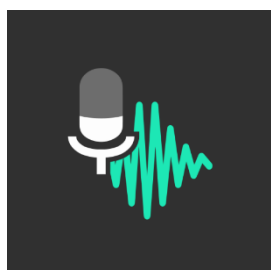
Στα παραπάνω θα αναφερθούμε περισσότερο στο κεφάλαιο της υλοποίησης.

Συνοπτικά όμως, θα ανιχνεύουμε θόρυβο με τη χρήση των παραπάνω φασματικών χαρακτηριστικών και ανάλογα με τις τιμές που θα παίρνουμε θα βαθμολογείται η ηχογράφηση από 1 έως 5. Η βαθμολογία 1 σημαίνει πολύ θορυβώδες, και 5 σημαίνει πάρα πολύ καθαρή ηχογράφηση. (Vrysis, Vryzas, Sidiropoulos, Avraam, & Dimoulas, n.d.).

## 2 Επισκόπηση του πεδίου

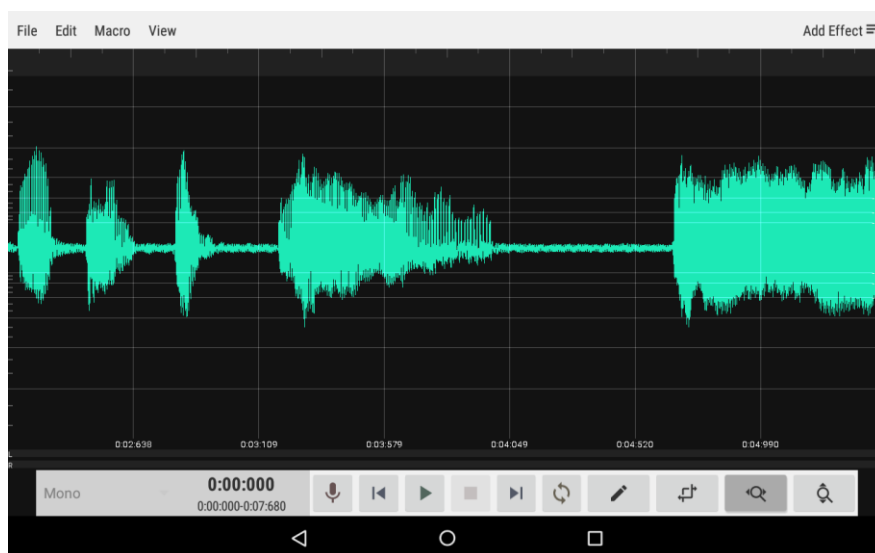
### 2.1 Εφαρμογές/Εργαλεία αξιολόγησης ηχογραφήσεων – μη αυτόματα

Στο λειτουργικό σύστημα Android υπάρχουν πολλές εφαρμογές που προσφέρουν τη δυνατότητα αξιολόγησης και επεξεργασίας ηχογραφήσεων με μη αυτόματο τρόπο. Αυτά τα εργαλεία απευθύνονται κυρίως σε μουσικούς, παραγωγούς, καθώς και σε επαγγελματίες του ήχου που επιθυμούν να έχουν πλήρη έλεγχο επί της ποιότητας και της επεξεργασίας των ηχογραφήσεών τους. Παρακάτω είναι μερικές από τις πιο δημοφιλείς εφαρμογές:



Εικόνα 6

**WaveEditor for Android™ Audio Recorder & Editor:** Η εφαρμογή WaveEditor πραγματοποιεί επεξεργασία ήχου, και υποστηρίζει πολλαπλούς τύπους αρχείων ήχου. Παρέχει εργαλεία όπως κύματα πολλαπλών παραθύρων, προσαρμογή της ταχύτητας αναπαραγωγής, και δυνατότητα ηχογράφησης με υψηλή ποιότητα.



Εικόνα 7: Διεπαφή της εφαρμογής WaveEditor for Android

Στην παραπάνω εικόνα βλέπουμε τη διεπαφή της εν λόγω εφαρμογής, όπου μας δείχνει την κυματομορφή του ηχητικού σήματος που φορτώθηκε ή καταγράφηκε και μας δίνει επιλογές για επεξεργασία, αναπαραγωγή, εφαρμογή εφέ και αποθήκευση.



Εικόνα 8

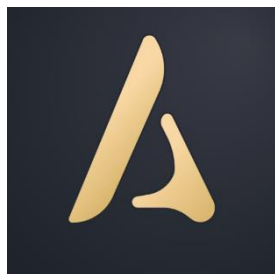
**FL Studio Mobile:** Αυτή η εφαρμογή είναι σχεδιασμένη για μουσική παραγωγή και σύνθεση. Προσφέρει πλήρη λειτουργικότητα DAW (Digital Audio Workstation) σε περιβάλλον κινητής συσκευής (smartphone ή tablet). Περιλαμβάνει δυνατότητες όπως σύνθεση ήχων, μίξη πολλαπλών καναλιών, και εφαρμογή εφέ.



Εικόνα 9: Διεπαφή της εφαρμογής FL Studiio Mobile

Όπως βλέπουμε στην παραπάνω εικόνα οι λειτουργίες και η διεπαφή είναι σχεδόν πανομοιότυπες με την εφαρμογή για σταθερό υπολογιστή. Τα διάφορα χειριστήρια και

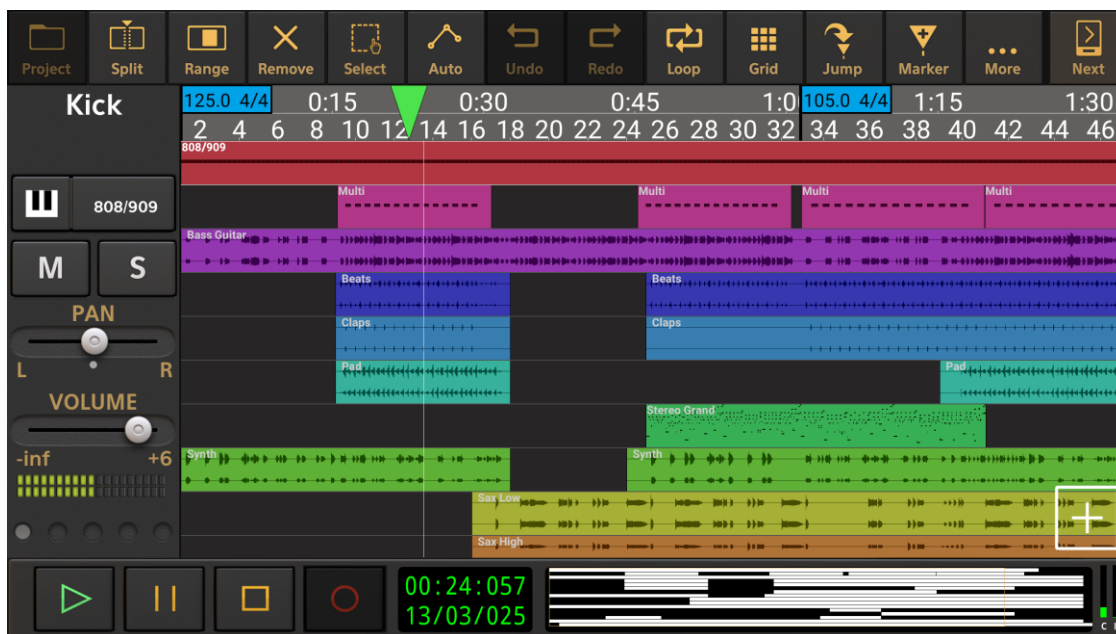
ενδείξεις δείχνουν σαφή προσανατολισμό προς τη χρήση από μουσικούς και για μουσική δημιουργία, κυρίως με συνθετικούς ήχους και MIDI.



Εικόνα 10

**Audio Evolution Mobile Studio:** Αυτή η εφαρμογή προσφέρει εξαιρετικά προηγμένες δυνατότητες επεξεργασίας και μίξης ήχου και είναι κατάλληλη για επαγγελματίες του ήχου και μουσικούς. Με πολυάριθμα εργαλεία και εφέ, αυτή η εφαρμογή είναι σχεδιασμένη για δημιουργία υψηλού επιπέδου μουσικής παραγωγής.

Οι παραπάνω εφαρμογές παρέχουν ένα ευρύ φάσμα δυνατοτήτων για την αξιολόγηση και επεξεργασία ηχογραφήσεων, από την βασική επεξεργασία έως την πλήρη μουσική παραγωγή, προσφέροντας έτσι στους χρήστες τη δυνατότητα να εργάζονται με ποιοτικό ήχο σε ένα πολύ ευέλικτο περιβάλλον.



Εικόνα 11: Διεπαφή της εφαρμογής Audio Evolution Mobile Studio

Στην παραπάνω εικόνα βλέπουμε την διεπαφή της εφαρμογής. Οι λειτουργίες που παρέχει είναι εξολοκλήρου για μουσική παραγωγή και επεξεργασία ηχητικού υλικού με εστίαση στη μίξη περισσότερο παρά στην σύνθεση. Παρέχει δυνατότητες αυτοματισμού καναλιών, λειτουργίες διόρθωσης τονικού ύψους (pitch correction), εφαρμογή εφέ και χρονικής επεξεργασίας/διόρθωσης.

Οι παραπάνω εφαρμογές αποτελούν δείγμα από το σύνολο των αντίστοιχων όμοιων εφαρμογών επεξεργασίας και παραγωγής ήχου/μουσικής και σε καμία περίπτωση δεν αποτελούν πλήρη λίστα των υπάρχουσών εφαρμογών. Σε όλες τις περιπτώσεις όμως ηχογραφούνται ή φορτώνονται αρχεία ηχογραφήσεων, όταν δεν πρόκειται για MIDI, και η αρχική ποιότητα της όποιας ηχογράφησης είναι σημαντικός παράγοντας για την επίτευξη του όποιου σκοπού αυτών των εφαρμογών και τις επεξεργασίες που δύναται να δεχθεί ένα αρχείο για τους σκοπούς αυτούς.

## **2.2 Εργαλεία/Ερευνες για την Αυτοματοποιημένη βελτίωση ηχογραφήσεων**

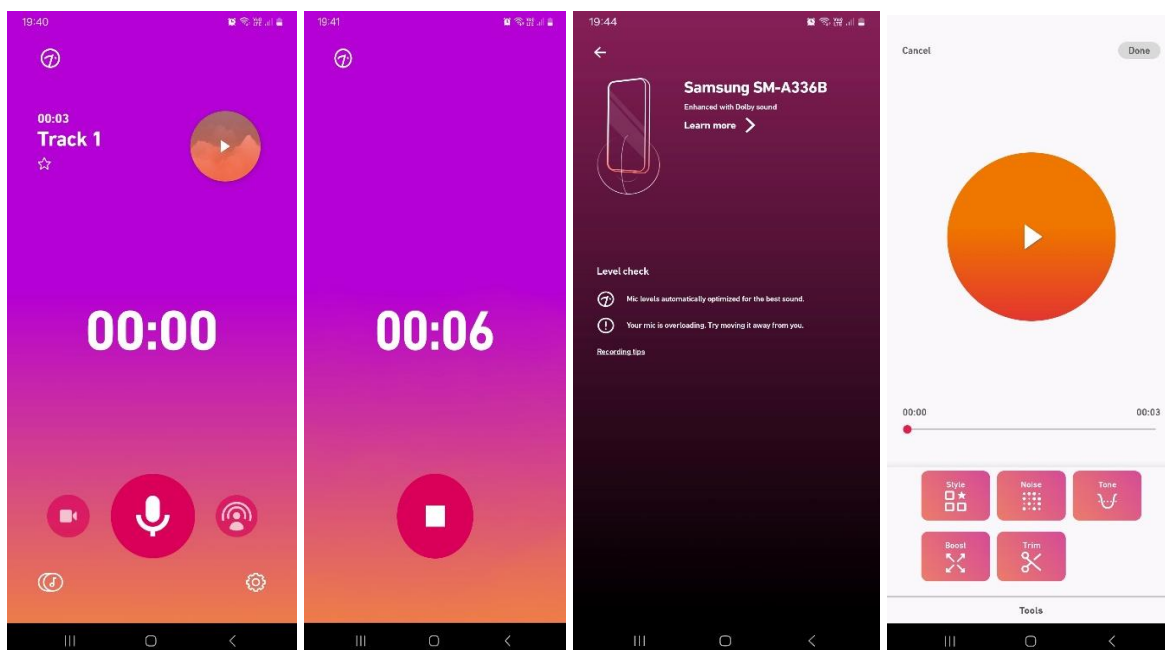
Στο λειτουργικό σύστημα Android υπάρχουν πολλές εφαρμογές που προσφέρουν τη δυνατότητα αυτόματης αξιολόγησης ηχογραφήσεων. Αυτά τα εργαλεία ενσωματώνουν σύγχρονες τεχνολογίες όπως την τεχνητή νοημοσύνη και τη μηχανική μάθηση για να αναλύουν, να επεξεργάζονται και να βελτιώνουν ηχογραφήσεις με ελάχιστη ή καθόλου ανθρώπινη παρέμβαση. Ακολουθούν δύο από τις κορυφαίες εφαρμογές για αυτόματη αξιολόγηση ηχογραφήσεων σε Android:



**Dolby On:** Είναι μια εφαρμογή σχεδιασμένη για την αυτόματη βελτίωση ηχογραφήσεων με χρήση προηγμένων αλγορίθμων της Dolby. Η εφαρμογή αναλύει τον ήχο και βελτιώνει αυτόματα την ποιότητα με βελτιστοποίηση έντασης και μείωση θορύβου.

**Εικόνα 12**

Όπως φαίνεται στις παρακάτω εικόνες οι λειτουργίες του περιλαμβάνουν πληθώρα επεξεργασιών για τη βελτίωση της ποιότητας μιας ηχογράφησης, μεταξύ των οποίων, αποθορυβοποίηση, μεγιστοποίηση έντασης/ακουστότητας (loudness), και ισοστάθμιση, οι οποίες πραγματοποιούνται αμέσως μόλις ο χρήστης πατήσει το πλήκτρο stop. Παρακάτω φαίνονται εικόνες από την λειτουργία της εφαρμογής και συμβουλές για βελτίωση της ηχογράφησης που εμφανίστηκαν όταν πατήσαμε το εικονίδιο επάνω αριστερά.



Εικόνα 13: Τέσσερα στάδια της λειτουργίας της εφαρμογής

Οι παραπάνω εικόνες αποτελούν στιγμιότυπα από τη λειτουργία της εφαρμογής όπως περιγράφηκε παραπάνω και η τελευταία εικόνα είναι η διεπαφής όταν επιλέγουμε μία ηχογράφηση, όπου μας εμφανίζεται ένας player, και με ενεργά εικονίδια από κάτω, διάφορες επεξεργασίες που μπορούν να πραγματοποιηθούν στο αρχείο, καθώς και δυνατότητες αρχειοθέτησης και διαχείρισης των αρχείων που δημιουργούνται από την εφαρμογή.

Αυτές οι εφαρμογές αξιοποιούν τις δυνατότητες της τεχνητής νοημοσύνης για να παρέχουν αξιόπιστες και αποδοτικές λύσεις για την αυτόματη αξιολόγηση ηχογραφήσεων, προσφέροντας έτσι εργαλεία που είναι ικανά να βελτιώσουν σημαντικά την ποιότητα του ήχου με ελάχιστη προσπάθεια από τον χρήστη.





Εικόνα 14

**Voice Recorder :** Η εφαρμογή αυτή αποτελεί μία τυπική εφαρμογή ηχογράφησης φωνής με δυνατότητα επεξεργασίας και αρχειοθέτησης των ηχογραφημένων αρχείων, επισημάνσης τμημάτων και λειτουργία απομαγνητοφώνησης. Δίνει τη δυνατότητα επιλογής σε τι ποιότητα θα γίνει η ηχογράφηση, με τρία προεπιλεγμένα σεντ ρυθμίσεων, για μέγιστη ποιότητα, για τον καλύτερο συμβιβασμό μεταξύ ποιότητας και εξοικονόμησης χώρου, και μεγιστοποίηση εξοικονόμησης χώρου.

Οι τρεις καταστάσεις λειτουργίας αφορούν το AAC codec και μόνο, αλλά μπορεί να γίνει καταγραφή και σε mp3 το οποίο και αυτό δίνεται η δυνατότητα να παραμετροποιηθεί αντίστοιχα. Εκτός από απωλεστικές, γράφει και σε μη απωλεστικές μορφές αρχείων, όπως flac αλλά και wav PCM σε ποιότητα μέχρι 44.1kHz, 16bit, mono. Η pro έκδοση δίνει τη δυνατότητα για stereo ηχογράφηση. Ενώ στις προχωρημένες ρυθμίσεις, δίνεται η δυνατότητα για επιλογή πηγής, όπως π.χ. το μικρόφωνο της κάμερας, το βασικό μικρόφωνο της συσκευής ή το μικρόφωνο απ' το Bluetooth. Δίνεται τέλος η δυνατότητα να εφαρμοστούν τρία εφέ/επεξεργασίες στο ηχογραφημένο σήμα για τη βελτίωση της ηχογράφησης, όπως διαγραφή τμημάτων της ηχογράφησης που περιέχουν σιωπή, αφαίρεση του θορύβου βάθους και ακύρωση ανακλάσεων.

Οι δύο εφαρμογές που παρουσιάζονται παραπάνω παρέχουν λειτουργίες με σκοπό την εξυπηρέτηση δύο διαφορετικών κοινών, πράγμα το οποίο φαίνεται και από τον τρόπο με τον οποίο προωθούνται.

Το Dolby On, προβάλλοντας τις ευκολίες και τα social χαρακτηριστικά της εφαρμογής, προωθείται περισσότερο για χρήση από μουσικούς και content creators οι οποίοι ως μη ειδικοί να μπορούν αυτόματα με ελάχιστες κινήσεις να έχουν τον καλύτερο δυνατό ήχο στην ηχογράφησή τους και μπορούν απευθείας να δημοσιεύσουν το υλικό τους.

Από την άλλη το Voice Recorder, φαίνεται να έχει περισσότερο στόχο χρήστες όπως δημοσιογράφους και φοιτητές/μαθητές, εστιάζοντας αμιγώς στην ηχογράφηση φωνής και στη λειτουργία απομαγνητοφώνησης.



### 3 Μεθοδολογία και υλοποίηση

Στο κεφάλαιο αυτό γίνεται μία εισαγωγή στις βασικές έννοιες και εργαλεία για που χρησιμοποιήθηκαν για τη δημιουργία της εφαρμογής. Στη συνέχεια παρουσιάζουμε τη συλλογιστική που ακολουθήσαμε για τη συγγραφή του κώδικα της εφαρμογής και τέλος, στιγμιότυπα από τη λειτουργία της εφαρμογής.

#### 3.1 Το λειτουργικό σύστημα φορητών συσκευών Android

Το Android είναι ένα λειτουργικό σύστημα βασισμένο σε Linux, το οποίο σχεδιάστηκε κυρίως για κινητές συσκευές με οθόνη αφής όπως τα smartphones και τα tablets. Αναπτύχθηκε αρχικά από την Android Inc., την οποία απέκτησε η Google το 2005. Η πρώτη εμπορική έκδοση του Android κυκλοφόρησε το 2008. Με την πάροδο των ετών, το Android έχει επεκταθεί για να υποστηρίξει ποικιλία άλλων τεχνολογιών και έχει γίνει το δημοφιλέστερο λειτουργικό σύστημα σε κινητές συσκευές σε παγκόσμιο επίπεδο (Android.com, 2023).

Η αρχιτεκτονική του Android επιτρέπει στους προγραμματιστές να χρησιμοποιούν την Java γλώσσα προγραμματισμού για τη δημιουργία εφαρμογών που τρέχουν στην ειδική εικονική μηχανή του (Dalvik Virtual Machine) ή στην Android Runtime (ART). Τα apps του Android διανέμονται κυρίως μέσω του GooglePlay Store, αν και υπάρχουν και άλλες πλατφόρμες διανομής εφαρμογών.

Εκτός από τις εφαρμογές, το Android υποστηρίζει ευρεία γκάμα συσκευών. Οι κατασκευαστές μπορούν να προσαρμόσουν το λειτουργικό σύστημα για να ταιριάζει στις δικές τους ανάγκες, πράγμα που έχει ως αποτέλεσμα μια μεγάλη ποικιλία συσκευών Android στην αγορά, από συσκευές υψηλών προδιαγραφών μέχρι πιο προσιτές επιλογές. Αυτή η ευελιξία έχει συμβάλει σημαντικά στην ευρεία διάδοση του συστήματος.

Η διαχείριση της ιδιωτικότητας και ασφάλειας στο Android είναι ένα σημαντικό θέμα, με την Google να παρέχει τακτικές ενημερώσεις και patches ασφαλείας. Επιπλέον, το Android υποστηρίζει μια σειρά από λειτουργίες όπως πολλαπλοί λογαριασμοί χρηστών, δυνατότητες προσβασιμότητας, και εξατομικευμένες ρυθμίσεις οθόνης.

Με την τεχνολογική εξέλιξη, το Android συνεχίζει να αναπτύσσεται και να ενσωματώνει νέες τεχνολογίες, όπως υποστήριξη για τεχνητή νοημοσύνη, επαυξημένη και εικονική πραγματικότητα, καθιστώντας το ένα ισχυρό εργαλείο για χρήστες και προγραμματιστές παγκοσμίως.

### **3.2 Η βιβλιοθήκη ψηφιακής επεξεργασίας σήματος JTransforms**

Η JTransforms είναι μια βιβλιοθήκη ανοιχτού κώδικα για ψηφιακή επεξεργασία σήματος (DSP), γραμμένη εξ ολοκλήρου σε Java. Αναπτύχθηκε για να παρέχει μια πολυνηματική υλοποίηση των πιο διαδεδομένων μετασχηματισμών Fourier, προσφέροντας σημαντική απόδοση και ευελιξία στις εφαρμογές που τη χρησιμοποιούν (GitHub, 2023).

Η βιβλιοθήκη JTransforms υποστηρίζει τέσσερις κύριους τύπους μετασχηματισμών: Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), Discrete Sine Transform (DST), Discrete Hartley Transform (DHT). Αυτοί οι μετασχηματισμοί είναι βασικοί για την ανάλυση και επεξεργασία σήματος, επιτρέποντας στους προγραμματιστές να εφαρμόζουν λειτουργίες όπως φιλτράρισμα, ανάλυση συχνοτήτων και άλλες επεξεργασίες σήματος. Η JTransforms είναι σχεδιασμένη να εκμεταλλεύεται την πολυνηματική επεξεργασία, βελτιώνοντας έτσι την απόδοση σε πολυπύρηννα συστήματα (GitHub, 2023).

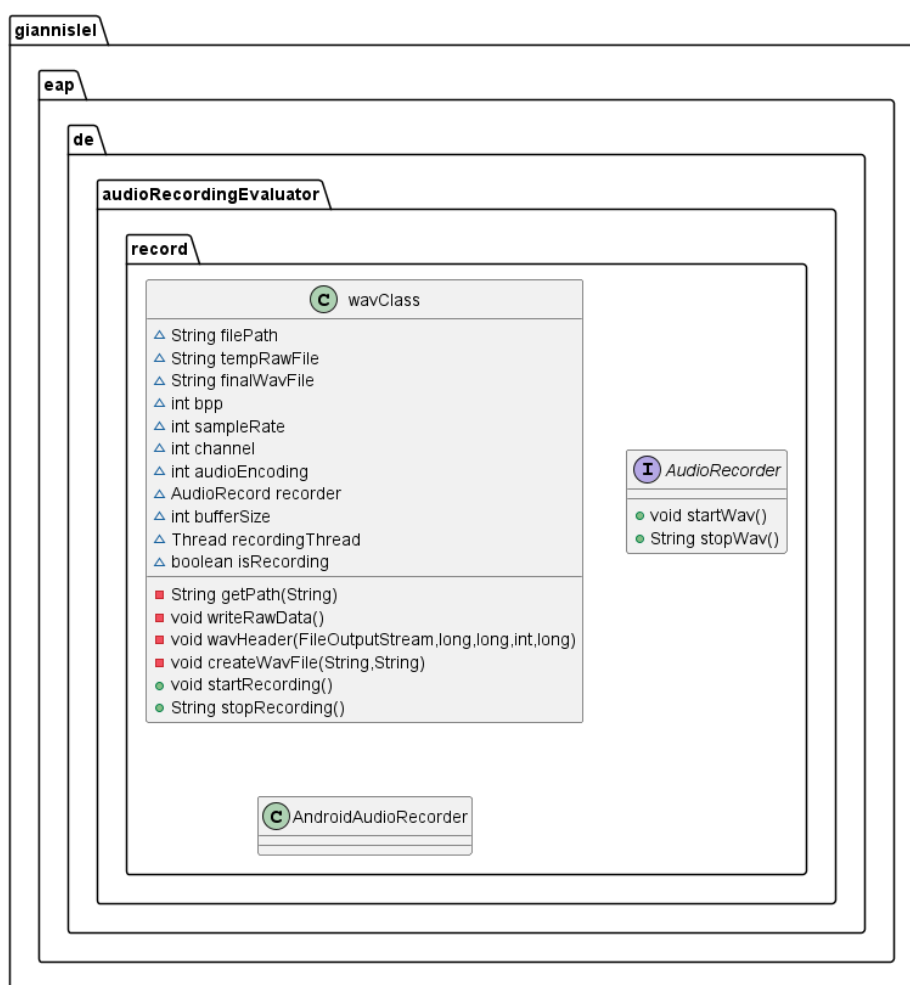
Η χρήση της βιβλιοθήκης είναι ευρέως διαδεδομένη σε διάφορους τομείς που απαιτούν ψηφιακή επεξεργασία σήματος. Εφαρμογές περιλαμβάνουν την ανάλυση ήχου, την επεξεργασία εικόνας, τις τηλεπικοινωνίες και πολλές άλλες επιστημονικές και μηχανολογικές εφαρμογές. Ένα άλλο χαρακτηριστικό της JTransforms είναι η συμβατότητά της με άλλες Java βιβλιοθήκες, όπως η Apache Commons Math, διευκολύνοντας έτσι την ενσωμάτωση σε μεγαλύτερα έργα λογισμικού (JorenSix/TarsosDSP, 2023).

Η ανάπτυξη της JTransforms συνεχίζεται με την κοινότητα ανοιχτού κώδικα να συμβάλλει στη βελτίωση και επέκταση των δυνατοτήτων της. Οι χρήστες μπορούν να βρουν τον κώδικα και την τεκμηρίωση στη σελίδα του GitHub του έργου, καθώς και παραδείγματα χρήσης και συγκρίσεις απόδοσης με άλλες βιβλιοθήκες (Aquila DSP, 2023).

Για τις ανάγκες των υπολογισμών μας, από τη βιβλιοθήκη αυτή χρησιμοποιήσαμε την κλάση FloatFFT\_1D. Για τον υπολογισμό του φάσματος ισχύος και άλλων χαρακτηριστικών του σήματος που καταγράψαμε χρησιμοποιήθηκαν οι μέθοδοι της κλάσης FFTUtils .

### 3.3 Καταγραφή PCM ηχητικών δεδομένων από το μικρόφωνο της κινητής συσκευής

Για την καταγραφή της ηχητικής πληροφορίας από το μικρόφωνο της συσκευής χρησιμοποιήθηκαν οι κλάσεις του πακέτου record. Στην παρακάτω εικόνα βλέπουμε τα διαγράμματα κλάσεων του πακέτου αυτού.



Εικόνα 15: Διάγραμμα κλάσης του πακέτου record

Από την κλάση `AndroidAudioRecorder` καλείται η μέθοδος `startWav` που εκκινεί την ηχογράφιση, καλώντας τη μέθοδο `startRecording` από την κλάση `wavClass`. Εντός της `wavClass`, αρχικά δημιουργείται ένα αρχείο για την προσωρινή αποθήκευση των δειγμάτων, για όσο γίνεται ηχογράφιση. Μόλις τελειώσει η ηχογράφιση καλείται η `stopwav` και τα δείγματα από το προσωρινό αρχείο μεταφέρονται στο τελικό αρχείο και στο ίδιο προστίθεται η RIFF κεφαλίδα (RIFF header) του wav αρχείου ώστε αυτό να αναγνωρίζεται ως τέτοιο, και η ίδια μέθοδος επιστρέφει τη διαδρομή του τελικού αρχείου στη μνήμη της συσκευής.

Εντός της `wavClass`, πρώτη καλείται η μέθοδος κατασκευαστή, με τη δημιουργία του αντικειμένου `recorderWav`, με όρισμα τη διαδρομή στο σύστημα αρχείων της συσκευής όπου θα δημιουργηθεί το αρχείο που θα περιέχει την ηχογράφιση, και υπολογίζει το ελάχιστο buffer που θα χρειαστεί για την καταγραφή μέσω της μεθόδου `getMinBufferSize` της κλάσης `AudioRecord`. Στη συνέχεια, όταν καλείται η μέθοδος `startRecording`, δημιουργείται ένα αντικείμενο της κλάσης `AudioRecord`, με ορίσματα, τη συσκευή καταγραφής, τη συχνότητα δειγματοληψίας, τον αριθμό των καναλιών της ηχογράφησης και το μέγεθος του buffer. Έπειτα γίνεται έλεγχος αν είναι διαθέσιμη η συσκευή καταγραφής, το μικρόφωνο σε αυτή την περίπτωση. Αφού διασφαλιστεί ότι η συσκευή είναι διαθέσιμη, καλείται η `startRecording` της `AudioRecord`, και θέτουμε την μεταβλητή `isRecording` στην τιμή `true`, ώστε να έχουμε μία ένδειξη ότι συμβαίνει καταγραφή. Δημιουργείται νήμα για την καταγραφή με όρισμα τη μέθοδο `writeRawData` για την καταγραφή των δεδομένων στο προσωρινό αρχείο με τη μορφή ενός πίνακα τύπου `byte`. Και εκκινείται το νήμα για την καταγραφή. Στη μέθοδο `writeRawData` ελέγχεται αν η ένδειξη `isRecording` είναι αληθής (ότι δεν έχει σταματήσει η ηχογράφιση) και αν επιτρέπεται η ανάγνωση των δεδομένων. Πραγματοποιείται καταγραφή των δεδομένων, μέσω της ροής `fileOutputStream`, στο προσωρινό αρχείο, μέχρι η ένδειξη ηχογράφησης να γίνει `false`, οπότε και κλείνει η ροή προς το προσωρινό αρχείο.

Όταν πλέον κληθεί η `stopRecording` και σταματήσει η καταγραφή των `byte` δεδομένων από το μικρόφωνο, το πρώτο που γίνεται εφόσον είναι όλα ομαλά και το αντικείμενο `recorder` υπάρχει, είναι να ανατεθεί η τιμή `false` στην μεταβλητή `isRecording`, ώστε να δειχθεί με κάποιο τρόπο ότι η ηχογράφιση σταμάτησε. Ελέγχεται ξανά η διαθεσιμότητα της συσκευής και καλείται η μέθοδος `stop()` για να σταματήσει η καταγραφή, με τη `release()`

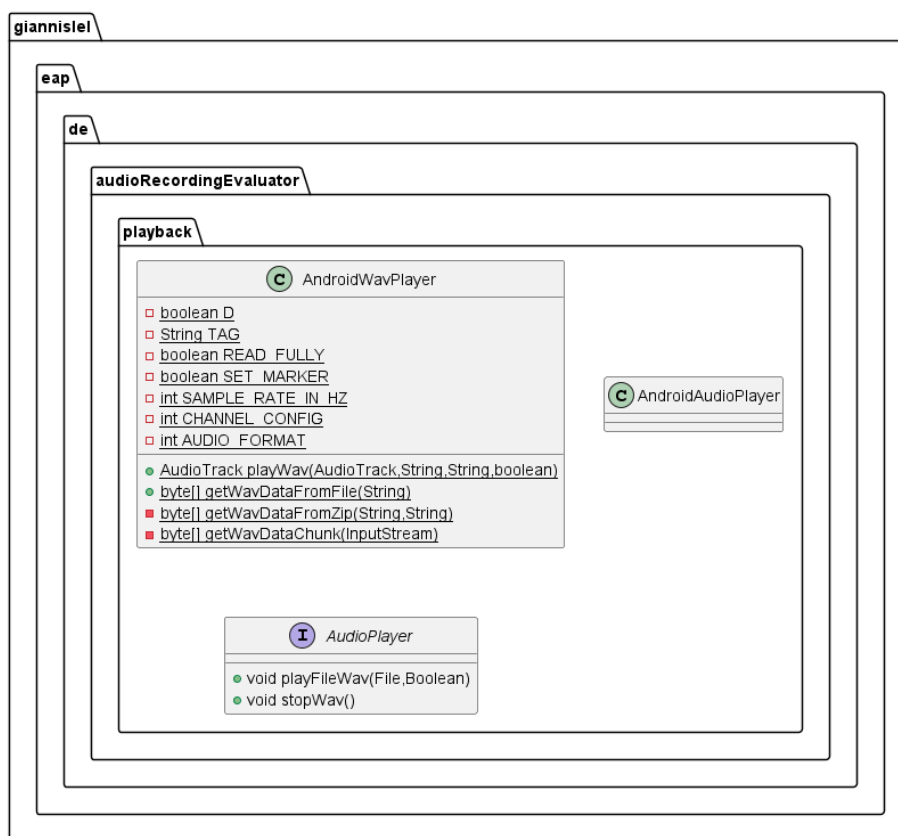
απελευθερώνονται οι πόροι που είχαν δεσμευτεί για την καταγραφή και το νήμα της καταγραφής γίνεται null. Τέλος, καλείται η createWavFile με ορίσματα, τη διαδρομή του προσωρινού αρχείου καθαρών δεδομένων, και αυτή του τελικού wav αρχείου, ώστε να γίνει η κατασκευή του τελικού αρχείου.

Εντός της createWavFile, δημιουργούνται δύο ροές δεδομένων, fileInputStream, από το προσωρινό αρχείο, και fileOutputStream, προς το τελικό. Στη συνέχεια δημιουργούνται μεταβλητές με τα διάφορα χαρακτηριστικά του αρχείου για να χρησιμοποιηθούν ως ορίσματα στην wavheader, ώστε η κεφαλίδα του αρχείου να έχει σωστές τιμές. Δημιουργείται πίνακας byte για την αποθήκευση των δεδομένων, δηλώνεται ο αριθμός των καναλιών, ο ρυθμός byte ανά δευτερόλεπτο, και τέλος το συνολικό μέγεθος της audio πληροφορίας, και το συνολικό μέγεθος των δεδομένων. Καλείται η μέθοδος wavHeader, η οποία με βάση τα παραπάνω, γράφει τα δεδομένα της κεφαλίδας στο τελικό αρχείο.

Στη συνέχεια πίσω στην createWavFile, γράφονται τα audio δεδομένα στο αρχείο και όταν ολοκληρωθεί η μεταφορά, κλείνουν οι ροές fileInputStream και fileOutputStream.

### 3.4 Αναπαραγωγή των αρχείων ηχογραφήσεων

Για την αναπαραγωγή των αρχείων που περιέχουν τις ηχογραφήσεις που έγιναν, χρησιμοποιήθηκαν οι κλάσεις του πακέτου playback. Στην παρακάτω εικόνα βλέπουμε τα διαγράμματα κλάσεων του πακέτου αυτού.



Εικόνα 16: Διάγραμμα κλάσης του πακέτου playback

Οι κλάσεις του πακέτου αυτού καλούνται όταν ο χρήστης πατήσει το πλήκτρο play ώστε να γίνει αναπαραγωγή του όποιου αρχείου έχει ηχογραφηθεί ή φορτωθεί.

Στην MainActivity, το αντικείμενο playerWav, καλεί την playFileWav με όρισμα το αντικείμενο audioFileWav, το οποίο είναι και το αντικείμενο που κρατάει τα δεδομένα της ηχογράφησης. Η playFileWav ανήκει στην κλάση AndroidAudioPlayer που έχει δύο μεθόδους, την playFileWav και την stopWav (η stopWav δεν είναι υλοποιημένη), καλεί την playWav με όρισμα την απόλυτη διαδρομή του αρχείου της ηχογράφησης, και επιστρέφει

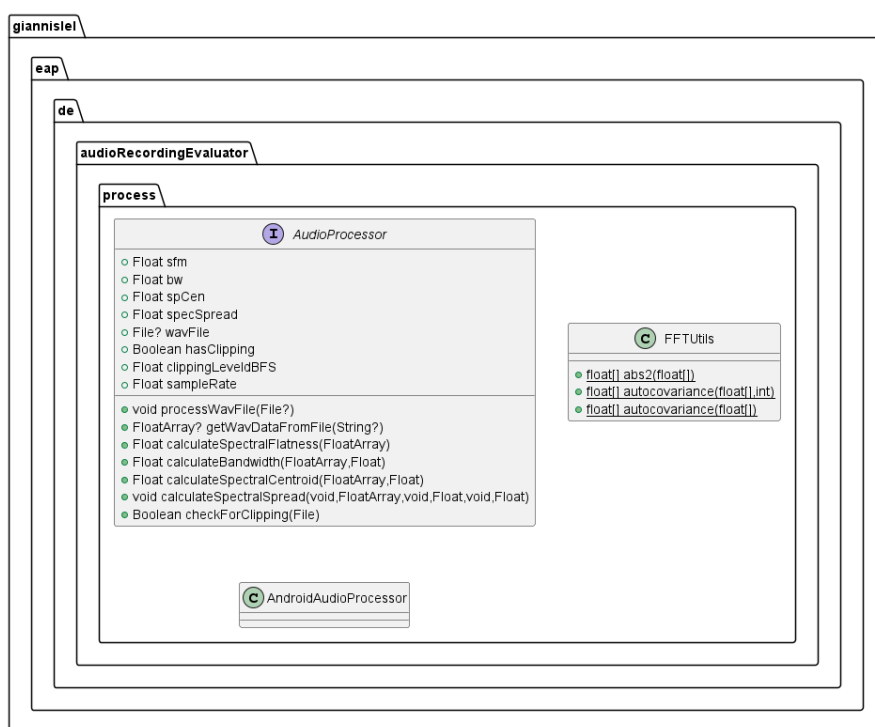
ένα αντικείμενο της κλάσης `AudioTrack`. Επιπλέον, η `playWav` ανήκει στην κλάση `AndroidWavPlayer`, και παίρνει ορίσματα που αφορούν αν τα δεδομένα προέρχονται από αντικείμενο `AudioTrack`, `ZipFile`, τη διαδρομή στην μνήμη του αρχείου, και μία ένδειξη για το αν τα δεδομένα προέρχονται από ροή δεδομένων.

Η `playWav`, αρχικά καλεί την `getWavDataFromFile` για να φορτώσει τα δεδομένα από το αρχείο σε μία προσωρινή μεταβλητή `buffer`. Τα δεδομένα αυτά χρησιμοποιούνται για τη δημιουργία ενός αντικειμένου `AudioTrack`. Το αντικείμενο αυτό δημιουργείται με όλα τα χαρακτηριστικά του αρχείου που φορτώνουμε για αναπαραγωγή, συχνότητα δειγματοληψίας, αριθμός καναλιών, `format` των δεδομένων (PCM 16 bit στην περίπτωση μας), το μήκος του πίνακα των δεδομένων. Μετά από διάφορους ελέγχους, ότι το αρχείο όντως υπάρχει και δεν είναι κενό, δημιουργείται ένα αντικείμενο `AudioTrack`, `audiotrack2`, με το οποίο γίνεται όλη η διαχείριση της καταγραφής των δεδομένων στο αντικείμενο που επιστρέφεται. Δημιουργείται ένα νήμα, εντός του οποίου γίνεται η ανάγνωση των audio δεδομένων.

Η `getWavDataFromFile`, που καλείται στην αρχή της `playWav`, για να πάρουμε τα PCM δεδομένα από το αρχείο. Εντός αυτής, φορτώνεται το αρχείο σε ένα αντικείμενο `FileInputStream`, `in`. Καλείται η `getWavDataChunk`, η οποία παρακάμπτει το κομμάτι της κεφαλίδας του αρχείου `wav`, και αποθηκεύει σε ένα πίνακα `byte` τα PCM δεδομένα. Τα δεδομένα αυτά αποθηκεύονται στη μεταβλητή `buffer`, και μεταφέρονται στην `playWav` που περιγράψαμε παραπάνω.

### 3.5 Επεξεργασία των ηχογραφημένων αρχείων

Για την επεξεργασία των αρχείων που περιέχουν τις ηχογραφήσεις που έγιναν, χρησιμοποιήθηκαν οι κλάσεις του πακέτου process. Στην παρακάτω εικόνα βλέπουμε τα διαγράμματα κλάσεων του πακέτου αυτού.



Εικόνα 17: Διάγραμμα κλάσης του πακέτου process

Όταν στην MainActivity, ο χρήστης πατήσει το πλήκτρο process, δημιουργείται το αντικείμενο processor της κλάσης AndroidAudioProcessor, και εκκινείται η διαδικασία επεξεργασίας για την εξαγωγή των χαρακτηριστικών του σήματος που χρησιμοποιούμε. Για να γίνει η επεξεργασία, καλείται η μέθοδος processWavFile με όρισμα ένα αντικείμενο τύπου File που περιέχει το αρχείο που δημιουργήσαμε για την ηχογράφηση. Εντός αυτής υπολογίζονται με τη σειρά, η ύπαρξη ψαλιδισμού κυματομορφής (clipping), το μέτρο φασματικής επιπεδότητας (spectral flatness measure), το εύρος ζώνης (bandwidth), το φασματικό κέντρο μάζας (spectral centroid), και η φασματική διασπορά (spectral spread).



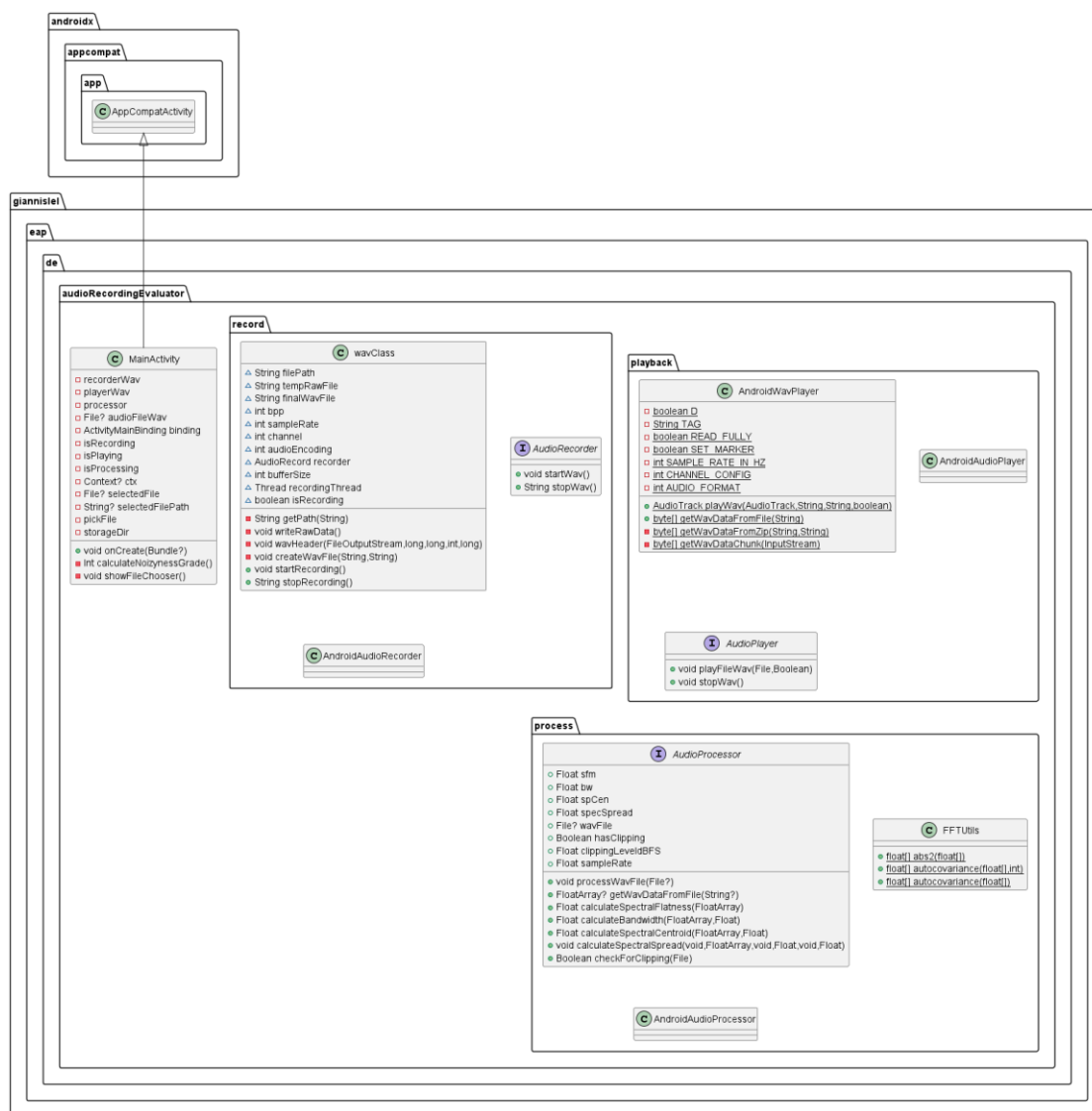
Για τον υπολογισμό των παραπάνω, αξιοποιείται η βιβλιοθήκη JTransforms, και η κλάση FloatFFT\_1D και η abs2 της κλάσης FFTUtils. Για την ανάκτηση των wav δεδομένων χρησιμοποιούμε μια παραλλαγή της getWavDataFromFile, η οποία δέχεται ένα όρισμα τύπου String για τη διαδρομή του αρχείου στη μνήμη της συσκευής και επιστρέφει τα audio δεδομένα υπό τη μορφή ενός FloatArray. Αυτή είναι η μορφή που δέχεται σαν είσοδο η FloatFFT\_1D. Καλούνται με τη σειρά οι μέθοδοι, checkForClipping, για την ανίχνευση clipping στο σήμα μας, calculateSpectralFlatness, για τον υπολογισμό του μέτρου φασματικής επιπεδότητας του σήματός μας, η calculateBandwidth για τον υπολογισμό του συχνοτικού εύρους του σήματός μας, η calculateSpectralCentroid, για τον υπολογισμό του φασματικού κέντρου μάζας του σήματός μας, και η calculateSpectralSpread, για τον υπολογισμό της φασματικής διασποράς του σήματός μας.

Η checkForClipping, διαφέρει από τις υπόλοιπες στο ότι κάνει τους υπολογισμούς σε δεδομένα τύπου byte, σε αντίθεση με τις υπόλοιπες μεθόδους που χρησιμοποιούν FloatArray λόγω της FloatFFT\_1D. Στις μεθόδους όπου χρησιμοποιείται η φασματική ισχύς για τους υπολογισμούς, καλείται η calcMagnitudeSpectrum, όπου αρχικά αντιγράφουμε την μεταβλητή floatArray σε μία μεταβλητή spectrum. Αυτή χρησιμοποιείται σαν όρισμα στην realForward της κλάσης FloadFFT\_1D για τον υπολογισμό του μονοδιάστατου διακριτού μετασχηματισμού φουριέ πάνω στο σήμα μας. Στη συνέχεια υπολογίζουμε την φασματική ισχύ του σήματος καλώντας την abs2 με όρισμα τα δεδομένα της μεταβλητής spectrum τα οποία περιέχουν πλέον το φάσμα του σήματος, το αποτέλεσμα αποθηκεύεται στην μεταβλητή magnitudeSpectrum και η οποία επιστρέφεται από τη μέθοδο.

### 3.6 Πλήρες Διάγραμμα Κλάσεων

Το παρακάτω διάγραμμα δείχνει όλες τις κλάσεις της εφαρμογής με όλες τις παραμέτρους και τις μεθόδους τους. Όπως φαίνεται, η λειτουργικότητα έχει χωριστεί σε τρία μέρη, αυτό της ηχογράφησης, αυτό της επεξεργασίας, και αυτό της αναπαραγωγής.

Η αλληλεπίδραση γίνεται εξολοκλήρου από την κλάση MainActivity, η οποία περιέχει το UI της εφαρμογής, το οποίο είναι ο διαμεσολαβητής μεταξύ του χρήστη και της όλης λειτουργικότητας που έχει ενσωματωθεί.



Εικόνα 18: Πλήρες διάγραμμα κλάσεων

Οι παραπάνω κλάσεις συνδέονται με τον εξής τρόπο, ο χρήστης αρχικά, είτε ηχογραφεί ένα αρχείο wav από το μικρόφωνο, είτε φορτώνει ένα αρχείο από την μνήμη, το οποίο έχει ηχογραφηθεί σε προγενέστερο χρόνο. Όταν πατηθεί το κουμπί Record, τότε καλούνται οι κλάσεις του πακέτου **record**, όταν πατηθεί το load, τότε καλούνται μέθοδοι από την κλάση MainActivity. Στη συνέχεια, εφόσον υπάρχει ένα wav αρχείο φορτωμένο, ο χρήστης μπορεί είτε να αναπαράξει το αρχείο ώστε να ακούσει την ηχογράφιση, οπότε και καλούνται οι κλάσεις του πακέτου **playback**, είτε να κάνει την επεξεργασία του αρχείου, οπότε καλούνται οι κλάσεις του πακέτου **process**, οι οποίες πραγματοποιούν τους υπολογισμούς των χαρακτηριστικών του σήματος μας, ώστε να αξιολογηθεί η ποιότητα της ηχογράφησης που είναι αποθηκευμένη στο αρχείο.

### 3.7 Μέθοδοι υπολογισμού παραμέτρων αξιολόγησης

Σε αυτή την ενότητα περιγράφουμε τις διάφορες μεθόδους υπολογισμού που χρησιμοποιήσαμε για να βρούμε τα χαρακτηριστικά του καταγεγραμμένου υλικού ώστε να βγάλουμε συμπεράσματα για το κατά πόσο θορυβώδης είναι η ηχογράφιση.

#### 3.7.1 Μέθοδος βαθμολόγησης ηχογράφησης

Για την παραγωγή της βαθμολογίας χρησιμοποιείται κλίμακα από 1 έως 5. Διαιρέσαμε το εύρος τιμών του SFM από 0 έως 1 σε πέντε ίσα τμήματα, δηλαδή ανά 0.2. Στην κλίμακά μας το 5 είναι η καλύτερη βαθμολογία και το 1 είναι η χαμηλότερη. Για κάθε 0.2 αύξηση της τιμής του SFM αφαιρείται ένας βαθμός από τη βαθμολογία. Το SFM αποτελεί το βασικό μέτρο που χρησιμοποιούμε για να κρίνουμε το κατά πόσο είναι θορυβώδες είναι το σήμα μας. Τα υπόλοιπα μεγέθη χρησιμοποιούνται βοηθητικά ώστε να βελτιώσουν τη βαθμολογία σε κάποιες περιπτώσεις, όπως για παράδειγμα στην περίπτωση που μπορεί να έχουμε αρκετό θόρυβο βάθους αλλά μικρό spectral spread και bandwidth, οπότε παρόλο που μπορεί το σήμα μας να έχει αρκετό θόρυβο, η καταληπτότητα είναι εν τέλει καλύτερη απ' ό,τι θα περιμέναμε σε αντίστοιχη περίπτωση.

Τα μεγέθη που χρησιμοποιούμε είναι η φασματική επιπεδότητα (SFM) που αναφέραμε προηγουμένως, το εύρος συχνοτήτων (bandwidth), στο τονικό κέντρο (spectral centroid), και η φασματική διασπορά (spectral spread). Τα βήματα που ακολουθούμε για να βγει η βαθμολογία είναι τα εξής:

Αρχικά υπολογίζουμε την φασματική επιπεδότητα ώστε να έχουμε μία πρώτη ένδειξη του κατά πόσο είναι θορυβώδες το σήμα μας αφού όσο πιο επίπεδο είναι το φάσμα του σήματός μας τόσο περισσότερο προσεγγίζει το λευκό θόρυβο.

Στη συνέχεια υπολογίζουμε το συχνοτικό εύρος του φάσματός μας. Σκοπός μας στον υπολογισμό αυτό είναι να μπορέσουμε να δούμε πόσο ευρύ είναι το φάσμα του σήματος που καταγράψαμε. Αν πάρουμε μικρές τιμές, σημαίνει ότι το σήμα μας είναι περιορισμένο στο εύρος αυτό οπότε και λιγότερο πιθανό να είναι θορυβώδες αυτό που καταγράφηκε (Brύσης et al., 2019).

Επόμενο βήμα είναι να υπολογίσουμε τη φασματική διασπορά του σήματός μας, αλλά για να γίνει αυτός ο υπολογισμός χρειάζεται να γνωρίζουμε και το τονικό κέντρο του σήματός μας. Για το λόγο αυτό, πρώτα υπολογίζουμε το τονικό κέντρο (spectral centroid) και στη συνέχεια με βάση αυτό υπολογίζουμε τη διασπορά γύρω από αυτό. Αυτό μας δείχνει ότι αν το φάσμα του σήματός μας είναι συγκεντρωμένο κοντά στο τονικό κέντρο, είναι πιο πιθανό να μην είναι θορυβώδες.

Με τον τρόπο αυτό μπορούμε να δούμε τέσσερα πράγματα για το σήμα μας. Πρώτον είναι θορυβώδες και σε τι βαθμό ποιο είναι το συχνοτικό του εύρος, εντός αυτού που βρίσκεται το τονικό του κέντρο, και τέλος, το κατά πόσο συγκεντρωμένη είναι η ενέργεια του σήματος στο κέντρο αυτό. Έτσι μπορούμε να κρίνουμε αν ένα σήμα είναι όντως θορυβώδες, για παράδειγμα, αν έχει μεγάλο μέτρο φασματικής επιπεδότητας, πάνω από 0.8, συχνοτικό εύρος σχεδόν όσο το ακουστικό φάσμα και συχνοτική διασπορά αντίστοιχα μεγάλη, τότε το σήμα μας είναι σίγουρα κυρίως θόρυβος. Ενώ αν αντίθετα είναι όλα τα παραπάνω μεγέθη πιο εστιασμένα, είναι πιθανό να έχει καλό λόγο σήματος προς θόρυβο το σήμα μας (Peeters, Giordano, Susini, Misdariis, & McAdams, 2011).

Ανεξάρτητα από τη βαθμολογία θορύβου, εξετάζεται και η ύπαρξη παραμόρφωσης ψαλιδισμού (clipping) στο ηχογραφημένο αρχείο. Για την πραγματοποίηση του ελέγχου αυτού, διαβάζουμε τα byte δεδομένα του αρχείου από το data chunk του wav αρχείου, μετατρέπουμε την τιμή σε dBFS και ελέγχουμε αν το σήμα μας υπερβαίνει κάποιο κατώφλι που ορίζουμε εμείς.

Ομοίως, για την αντίθετη περίπτωση, δίνεται η δυνατότητα να ελέγξουμε αν υπάρχει κάποιο πρόβλημα στο μικρόφωνο της συσκευής ή στη διασύνδεση κάποιου μικροφώνου που θα μπορούσε να συνδεθεί στη συσκευή μας. Αυτό μπορούμε να το πετύχουμε, ελέγχοντας, αρχικά, για σιωπή στο σήμα μας, στάθμη σήματος δηλαδή μικρότερη από -80dBFS, ή και για σύντομα διαστήματα τέτοιας σιωπής. Με τον τρόπο αυτό μπορούμε να ελέγξουμε, στην πρώτη περίπτωση, για παράδειγμα, αν παίρνουμε σήμα από το μικρόφωνο, είτε το ενσωματωμένο, είτε το εξωτερικό. Ενώ στην δεύτερη περίπτωση, αν η σύνδεση έχει κάποιο πρόβλημα και κάνει πολύ σύντομες διακοπές.

### 3.7.2 Ψαλιδισμός Κυματομορφής (Clipping)

Ο ψαλιδισμός κυματομορφής (waveform clipping) είναι μια μορφή παραμόρφωσης που συμβαίνει όταν το σήμα ήχου υπερβαίνει την μέγιστη επιτρεπτή τιμή του συστήματος μετατροπής αναλογικού σήματος σε ψηφιακό (ADC). Αυτό οδηγεί στην αποκοπή των κορυφών του σήματος, με αποτέλεσμα μια επίπεδη κορυφή στην κυματομορφή και σημαντική απώλεια ποιότητας του ήχου. Στον ψηφιακό ήχο, ο ψαλιδισμός συμβαίνει όταν το σήμα υπερβαίνει τα 0 dBFS, με αποτέλεσμα μια απότομη και δυσάρεστη παραμόρφωση (Unison, 2021).

Το φαινόμενο του ψαλιδισμού, στα πλαίσια ενός ψηφιακού συστήματος όπως ένα λογισμικό Η/Υ, μπορεί να γίνει αντιληπτό είτε μέσω οπτικών ή και ακουστικών ενδείξεων. Οπτικά, ένας ψαλιδισμένος ήχος μπορεί να αναγνωριστεί από τις επίπεδες κορυφές στην κυματομορφή όταν εμφανίζεται σε ένα πρόγραμμα επεξεργασίας ήχου. Πολλά ψηφιακά συστήματα επεξεργασίας ήχου παρέχουν ενδείξεις ψαλιδισμού (clipping) στα μετρητικά

τους, τα οποία συνήθως ανάβουν κόκκινα όταν ο ήχος υπερβαίνει το επιτρεπτό όριο (Hansen, Stauffer, & Xia, 2018).

Ακουστικά, ο ψαλιδισμός συχνά εκδηλώνεται με τη μορφή σκληρής, τραχιάς παραμόρφωσης, ειδικά στις υψηλές συχνότητες, προκαλώντας έναν δυσάρεστο και μεταλλικό ήχο. Μπορεί επίσης να συνοδεύεται από "κρότους" και "τρίξιμο" καθώς και σημαντική απώλεια καθαρότητας του σήματος (Unison, 2021).

Η πρόληψη του ψαλιδισμού περιλαμβάνει την προσεκτική ρύθμιση των επιπέδων εισόδου κατά την εγγραφή και την επεξεργασία ήχου, καθώς και τη χρήση δυναμικών επεξεργαστών όπως οι limiters για την αποφυγή υπερβολικών κορυφών στο σήμα (Unison, 2021).

### 3.7.3 Φασματική επιπεδότητα

Η φασματική επιπεδότητα (spectral flatness measure) είναι ένα μέτρο που χρησιμοποιείται στην επεξεργασία σήματος για να χαρακτηρίσει το φάσμα ενός ήχου. Συγκεκριμένα, η φασματική επιπεδότητα δείχνει πόσο "επίπεδο" ή "ομαλό" είναι το φάσμα ενός σήματος, και παρέχει έναν τρόπο για να ποσοτικοποιηθεί πόσο θορυβώδης είναι ένας ήχος σε σύγκριση με έναν πιο τονικό ή ηχηρό ήχο.

Ο υπολογισμός της φασματικής επιπεδότητας βασίζεται στην αναλογία της γεωμετρικής μέσης τιμής των πλατών του φάσματος προς την αριθμητική μέση τιμή των ίδιων πλατών. Ο μαθηματικός τύπος για τον υπολογισμό της φασματικής επιπεδότητας είναι ο εξής:

$$\text{SFM} = \frac{(\prod_{i=1}^N x_i)^{1/N}}{\frac{1}{N} \sum_{i=1}^N x_i}$$

όπου  $x_i$  είναι τα πλάτη του φάσματος και  $N$  ο αριθμός των στοιχείων στο φάσμα.

Η φασματική επιπεδότητα χρησιμοποιείται ευρέως στην ανάλυση και αναγνώριση ήχων, ιδιαίτερα στη μουσική και στην επεξεργασία λόγου. Ένας υψηλός δείκτης φασματικής επιπεδότητας υποδεικνύει ένα θορυβώδες σήμα, ενώ ένας χαμηλός δείκτης υποδηλώνει έναν τονικό ήχο με πιο καθαρό φάσμα. Αυτή η μέτρηση είναι χρήσιμη σε εφαρμογές όπως η

ανίχνευση ομιλίας, η ταξινόμηση μουσικών οργάνων και η ανάλυση ηχητικών σημάτων (Crystal Instruments, 2021; Keysight, 2021; MathWorks, 2021).

### 3.7.4 Συχνотικό Εύρος (Bandwidth)

Το συχνотικό εύρος ενός σήματος ορίζεται ως η διαφορά μεταξύ των ανώτερων και κατώτερων συχνοτήτων που συνθέτουν το σήμα. Αυτό το εύρος μετρίεται σε Hertz (Hz) και προσδιορίζει το φάσμα των συχνοτήτων που περιέχουν σημαντική ενέργεια του σήματος (Toppr, 2021).

Ο υπολογισμός του συχνотικού εύρους μπορεί να πραγματοποιηθεί λαμβάνοντας υπόψη τις ανώτερες και κατώτερες συχνότητες του σήματος. Η εξίσωση για τον υπολογισμό του εύρους είναι:

$$\text{Bandwidth} = f_H - f_L$$

Όπου  $f_H$  είναι η ανώτερη συχνότητα και  $f_L$  είναι η κατώτερη συχνότητα του σήματος (Test and Measurement Tips, 2021). Για παράδειγμα, αν ένα σήμα έχει συχνότητες από 635 MHz έως 7000 MHz, το συχνотικό εύρος του θα είναι 6365 MHz (Toppr, 2021).

Το συχνотικό εύρος είναι κρίσιμο για την ανάλυση και επεξεργασία ηχητικών σημάτων, καθώς καθορίζει το φάσμα των συχνοτήτων που περιέχουν την ενέργεια του σήματος. Στην επεξεργασία σήματος, το εύρος συχνοτήτων βοηθά στον προσδιορισμό της ποιότητας του ήχου και στη βελτιστοποίηση συστημάτων μετάδοσης ή επεξεργασίας ήχου (All About Circuits, 2021). Για παράδειγμα, στην ψηφιακή επεξεργασία σήματος, η εύρεση του συχνотικού εύρους είναι απαραίτητη για τη σωστή δειγματοληψία και την αποφυγή φαινομένων όπως το aliasing, όπως ορίζεται από το θεώρημα δειγματοληψίας Nyquist-Shannon (Siemens, 2021).

### 3.7.5 Φασματικό Κέντρο Μάζας (Spectral centroid)

Το φασματικό κέντρο μάζας ή φασματικό κεντροειδές (spectral centroid) είναι ένα σημαντικό μέτρο που χρησιμοποιείται για να περιγράψει τη "φωτεινότητα" ή τη "λαμπρότητα" ενός ηχητικού σήματος. Αντιπροσωπεύει, δηλαδή, τη συχνότητα γύρω από

την οποία συγκεντρώνεται η μάζα του φάσματος, δηλαδή, πόσο και πού είναι κεντραρισμένο το φάσμα.

Ο υπολογισμός του φασματικού κέντρου μάζας γίνεται λαμβάνοντας τον σταθμισμένο μέσο όρο των συχνοτήτων, με τα βάρη να είναι τα πλάτη των συχνοτήτων αυτών. Ο μαθηματικός τύπος είναι ο εξής:

$$SC = \frac{\sum_{i=1}^N f_i A_i}{\sum_{i=1}^N A_i}$$

όπου  $f_i$  είναι η συχνότητα  $A_i$  είναι το πλάτος της συχνότητας  $f_i$  (Peeters, 2004).

Η χρήση του φασματικού κέντρου μάζας είναι ιδιαίτερα σημαντική στη μουσική ανάλυση και στην αναγνώριση ομιλίας. Στη μουσική, το φασματικό κεντροειδές μπορεί να βοηθήσει στη διάκριση των μουσικών οργάνων και των μουσικών ειδών, καθώς διαφορετικά όργανα και είδη έχουν διαφορετικά φασματικά χαρακτηριστικά (Agostini, Longari, & Pollastri, 2003). Στην αναγνώριση ομιλίας, το φασματικό κεντροειδές χρησιμοποιείται για την αναγνώριση της "φωτεινότητας" της φωνής, βοηθώντας στην ταυτοποίηση ομιλητών και στην ανίχνευση αλλοιώσεων ήχου (Chen et al., 2004).

### 3.7.6 Spectral spread

Η φασματική διασπορά είναι ένα μέτρο που χρησιμοποιείται για να περιγράψει τη διασπορά ή την κατανομή της ενέργειας ενός ηχητικού σήματος γύρω από το φασματικό κεντροειδές. Συγκεκριμένα, η φασματική διασπορά υπολογίζεται ως η τετραγωνική ρίζα του σταθμισμένου μέσου όρου των τετραγωνικών αποστάσεων των συχνοτήτων από το φασματικό κεντροειδές. Ο τύπος για τη φασματική διασπορά είναι:

$$SS = \sqrt{\frac{\sum_{i=1}^N (f_i - SC)^2 A_i}{\sum_{i=1}^N A_i}}$$

Όπου  $f_i$ , είναι η συχνότητα,  $A_i$ , είναι το πλάτος της συχνότητας  $f_i$ , και  $SC$  είναι το φασματικό κέντρο μάζας (spectral centroid) (Peeters, 2004).



Η φασματική διασπορά παίρνει τιμές σε Hz υποδεικνύοντας το συχνотικό εύρος γύρω από το οποίο είναι κατανεμημένη η ενέργεια του σήματος γύρω από το φασματικό κεντροειδές. Χρησιμοποιείται συχνά για την αξιολόγηση της ποιότητας του ήχου και της ηχητικής υφής. Ένα σήμα με υψηλή φασματική διασπορά έχει μεγαλύτερη ποικιλία συχνοτήτων γύρω από το φασματικό κεντροειδές, ενώ ένα σήμα με χαμηλή φασματική διασπορά έχει συχνότητες συγκεντρωμένες γύρω από το κεντροειδές (Jiang et al., 2002). Αυτό το μέτρο είναι χρήσιμο για την ανάλυση μουσικών οργάνων και την αναγνώριση ηχητικών περιβαλλόντων (Agostini, Longari, & Pollastri, 2003).

### **3.8 Δημιουργία της Εφαρμογής**

#### **3.8.1 Android Studio**

Το Android Studio είναι το επίσημο ολοκληρωμένο περιβάλλον ανάπτυξης (IDE) της Google για την ανάπτυξη εφαρμογών Android. Παρουσιάστηκε το 2013 ως διάδοχος του Eclipse με το Android Development Tools (ADT) και έκτοτε έχει γίνει το πρότυπο εργαλείο για προγραμματιστές Android (Google, 2013).

Το Android Studio βασίζεται στην πλατφόρμα IntelliJ IDEA της JetBrains, προσφέροντας ένα ισχυρό σύνολο εργαλείων που επιτρέπουν στους προγραμματιστές να γράφουν, να δοκιμάζουν και να εντοπίζουν σφάλματα στις εφαρμογές τους. Ένα από τα πιο σημαντικά χαρακτηριστικά του Android Studio είναι το υποσύστημα Gradle για την κατασκευή έργων, το οποίο αυτοματοποιεί πολλές από τις διαδικασίες ανάπτυξης και επιτρέπει τη διαχείριση εξαρτήσεων και τη δημιουργία πολλαπλών παραλλαγών εφαρμογών (Toomey, 2014).

Το Android Studio προσφέρει επίσης έναν ενσωματωμένο εικονικό μηχανισμό (emulator) που επιτρέπει στους προγραμματιστές να δοκιμάζουν τις εφαρμογές τους σε μια ποικιλία συσκευών και διαμορφώσεων χωρίς να απαιτείται φυσική συσκευή. Επιπλέον, περιλαμβάνει εργαλεία ανάλυσης απόδοσης, όπως το Android Profiler, το οποίο παρέχει λεπτομερείς πληροφορίες για τη χρήση της CPU, της μνήμης και του δικτύου από την εφαρμογή (Shan, 2018).

Ένα άλλο σημαντικό χαρακτηριστικό του Android Studio είναι η υποστήριξη για την ανάπτυξη εφαρμογών με τη χρήση της γλώσσας Kotlin, η οποία αναγνωρίστηκε επίσημα από την Google ως γλώσσα ανάπτυξης για το Android το 2017. Η ενσωμάτωση της Kotlin έχει διευκολύνει τους προγραμματιστές να γράφουν πιο σαφή και συντηρήσιμο κώδικα (Borbely, 2018).

Το Android Studio συνεχίζει να εξελίσσεται με κάθε νέα έκδοση, προσφέροντας νέα εργαλεία και χαρακτηριστικά που βελτιώνουν την παραγωγικότητα των προγραμματιστών και την ποιότητα των εφαρμογών Android (Google, 2020).

### 3.8.2 Σχεδίαση διεπαφής χρήστη με τη χρήση των Views

Η σχεδίαση διεπαφής χρήστη (UI) στο λειτουργικό σύστημα Android βασίζεται στη χρήση των Views και των ViewGroups, τα οποία αποτελούν τα βασικά δομικά στοιχεία για τη δημιουργία διαδραστικών εφαρμογών. Τα Views είναι τα οπτικά στοιχεία που μπορεί να δει και να αλληλεπιδράσει ο χρήστης, όπως τα κουμπιά (Button), τα πεδία κειμένου (TextView), και οι εικόνες (ImageView) (Android Developers, 2023).

Για τη δημιουργία μιας διάταξης, χρησιμοποιούνται συνήθως αρχεία XML που περιγράφουν την ιεραρχία των Views και των ViewGroups. Αυτή η προσέγγιση διαχωρίζει τη λογική της εφαρμογής από την παρουσίαση, διευκολύνοντας την επαναχρησιμοποίηση και τη συντήρηση του κώδικα (C# Corner, 2023). Η πιο κοινή δομή περιλαμβάνει ViewGroups όπως το LinearLayout και το ConstraintLayout, που καθορίζουν τη διάταξη και τη θέση των Views μέσα στην οθόνη.

Το ConstraintLayout είναι ιδιαίτερα δημοφιλές για τη δημιουργία responsive και adaptive διεπαφών. Επιτρέπει τον καθορισμό σχέσεων μεταξύ των Views, διευκολύνοντας την προσαρμογή της διάταξης σε διαφορετικά μεγέθη και προσανατολισμούς οθόνης. Χρησιμοποιώντας το εργαλείο Layout Editor στο Android Studio, οι προγραμματιστές μπορούν να σύρουν και να αποθέτουν στοιχεία UI, να εφαρμόζουν constraints και να ρυθμίζουν τις ιδιότητες των Views χωρίς να χρειάζεται να επεξεργάζονται XML αρχεία με το χέρι (Android Developers, 2023).

Για παράδειγμα, το παρακάτω XML απόσπασμα δείχνει τη χρήση του ConstraintLayout για τη διάταξη ενός TextView και ενός Button:

```
<ConstraintLayout xmlns:android="http://schemas.android.com/apk/res/android"
```

```
    android:layout_width="match_parent"
```

```
    android:layout_height="match_parent">
```

```
    <TextView
```

```
        android:id="@+id/textView"
```

```
        android:layout_width="wrap_content"
```

```
        android:layout_height="wrap_content"
```

```
        android:text="Hello, World!"
```

```
        app:layout_constraintTop_toTopOf="parent"
```

```
        app:layout_constraintStart_toStartOf="parent"/>
```

```
    <Button
```

```
        android:id="@+id/button"
```

```
        android:layout_width="wrap_content"
```

```
        android:layout_height="wrap_content"
```

```
        android:text="Click Me"
```

```
        app:layout_constraintTop_toBottomOf="@id/textView"
```

```
        app:layout_constraintStart_toStartOf="parent"/>
```

```
</ConstraintLayout>
```

Αυτό το παράδειγμα δείχνει πώς τα Views μπορούν να τοποθετηθούν σε σχέση με το γονικό ViewGroup και μεταξύ τους, χρησιμοποιώντας constraints.

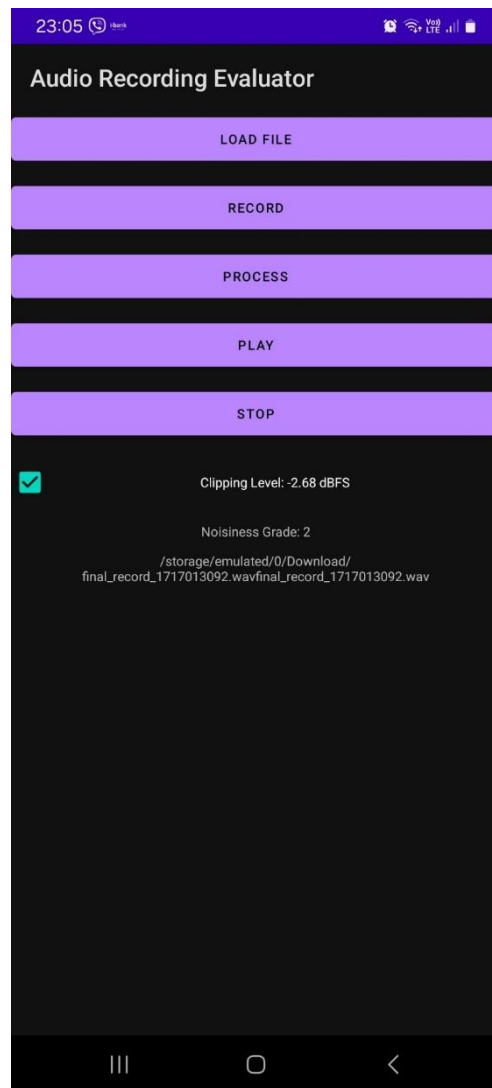
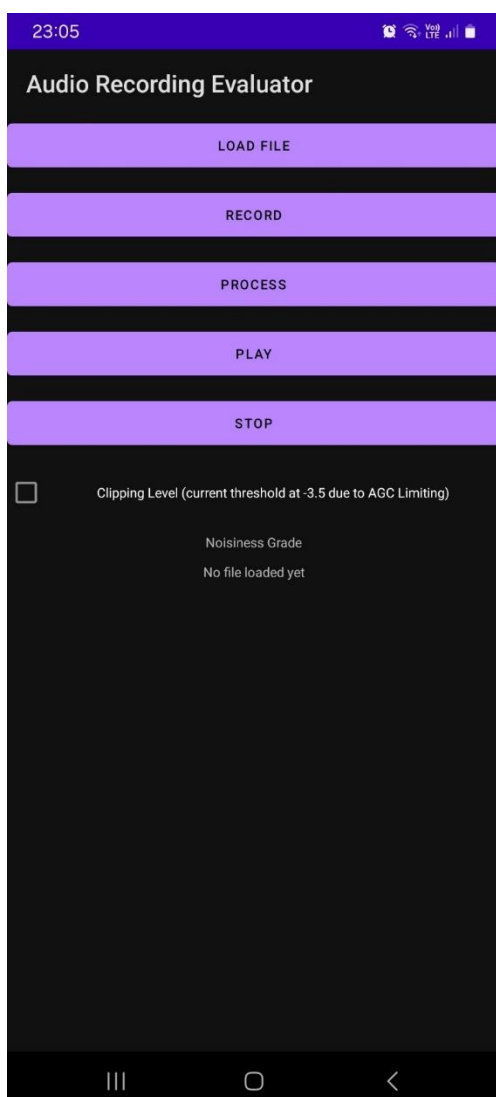
Η προσαρμοστικότητα των διατάξεων είναι κρίσιμη για την υποστήριξη ποικιλίας συσκευών με διαφορετικά μεγέθη οθόνης. Χρησιμοποιώντας παραλλαγές διατάξεων (layout variants), οι προγραμματιστές μπορούν να δημιουργήσουν ξεχωριστές διατάξεις για διαφορετικές καταστάσεις οθόνης, όπως η κατακόρυφη και η οριζόντια διάταξη, διασφαλίζοντας μια συνεπή εμπειρία χρήστη (Android Developers, 2023).

Υπάρχουν πλέον πιο εξελιγμένα μοντέλα ανάπτυξης διεπαφής χρήστη, όπως το Jetpack Compose, αλλά εμείς στην εφαρμογή μας, για λόγους απλότητας και ευκολίας στην υλοποίηση, χρησιμοποιήσαμε το View Model.

### 3.8.3 Γραφική διεπαφή χρήστη

Όπως φαίνεται στην εικόνα 5, η γραφική διεπαφή χρήστη αποτελείται από ένα πεδίο κειμένου, για τον τίτλο της εφαρμογής, 5 κουμπιά, τα οποία αφορούν τις λειτουργίες φόρτωσης, ηχογράφησης, επεξεργασίας, αναπαραγωγής, και λήξης είτε της αναπαραγωγής είτε της ηχογράφησης.

Ακολουθούν οι ενδείξεις που παράγονται από τα φασματικά χαρακτηριστικά του ηχογραφημένου ηχητικού υλικού. Πρώτο είναι η ένδειξη ψαλιδισμού, υλοποιημένη από ένα γραφικό στοιχείο checkbox, με το κείμενό του. Όταν υπάρχει ψαλιδισμός, το checkbox φαίνεται επιλεγμένο, και στο κείμενό του αναγράφεται η ένδειξη “Clipping Level: --- dBFS” η οποία υποδεικνύει τη μέγιστη στάθμη που έφτασε το σήμα μας μετά το κατώφλι που θέσαμε. Ακολουθεί η ένδειξη για το πόσο θόρυβο περιέχει η ηχογράφησή μας. Η ένδειξη αυτή παίρνει τιμές από 1 έως 5 με 1 την χειρότερη βαθμολογία και 5 την καλύτερη. Και Τέλος, ακολουθεί ένα πεδίο κειμένου με τη διαδρομή και το όνομα του αρχείου της ηχογράφησης στη μνήμη της συσκευής.



Εικόνα 19: (αριστερά) Γραφική διεπαφή χρήστη της εφαρμογής μας μόλις αυτή έχει ανοίξει  
(δεξιά) Γραφική διεπαφή χρήστη της εφαρμογής μας αμέσως μετά την ολοκλήρωση των επεξεργασιών  
του σήματος.

Τέλος, τα γραφικά της εφαρμογής περιλαμβάνουν και το εικονίδιο της εφαρμογής όπως αυτό φαίνεται στην παρακάτω εικόνα



**Εικόνα 20: Εικονίδιο Εφαρμογής Audio Recording Evaluator**

Στην παρακάτω εικόνα παρατίθεται QR code που περιέχει σύνδεσμο προς το Play Store της Google απ' όπου μπορεί να γίνει μεταφόρτωση και εγκατάσταση της εφαρμογής σε συμβατές συσκευές.



**Εικόνα 21: QR code με τον σύνδεσμο της εφαρμογής στο ηλεκτρονικό κατάστημα GooglePlay της Google**

## 4 Δοκιμές ευχρηστίας της εφαρμογής

Στο κεφάλαιο αυτό παρουσιάζουμε το θεωρητικό υπόβαθρο και τις μεθόδους που ακολουθήσαμε για να αξιολογήσουμε την εφαρμογή μας ως προς την ευχρηστία της. Παρουσιάζεται το ερωτηματολόγιο που απάντησαν οι χρήστες καθώς και λειτουργικός σύνδεσμος για τη φόρμα αξιολόγησης

### 4.1 Ευχρηστία

Η ευχρηστία είναι ένας θεμελιώδης παράγοντας που καθορίζει την ποιότητα και την αποδοχή μιας εφαρμογής από τους χρήστες. Ο όρος ευχρηστία αναφέρεται στον βαθμό στον οποίο μια εφαρμογή μπορεί να χρησιμοποιηθεί εύκολα και αποδοτικά από τους χρήστες για να επιτύχουν τους στόχους τους σε ένα συγκεκριμένο πλαίσιο χρήσης (Nielsen, 1993). Η ευχρηστία περιλαμβάνει διάφορες διαστάσεις, όπως η αποτελεσματικότητα, η αποδοτικότητα και η ικανοποίηση των χρηστών. Με καλά σχεδιασμένη ευχρηστία, μειώνεται η καμπύλη μάθησης, αυξάνεται η ικανοποίηση των χρηστών και βελτιώνεται η παραγωγικότητα. Στο πλαίσιο των εφαρμογών Android, η ευχρηστία αποτελεί κρίσιμο στοιχείο για την επιτυχία μιας εφαρμογής. Η πλατφόρμα Android είναι μία από τις πιο δημοφιλείς πλατφόρμες για κινητές συσκευές, με εκατομμύρια χρήστες παγκοσμίως. Επομένως, η ανάπτυξη εφαρμογών που είναι εύχρηστες και ευχάριστες στη χρήση μπορεί να καθορίσει τη διαφορά μεταξύ μιας επιτυχημένης και μιας αποτυχημένης εφαρμογής.

### 4.2 Αρχές Ευχρηστίας σε Εφαρμογές Android

**Απλότητα και Καθαρότητα Σχεδίασης:** Οι εφαρμογές Android πρέπει να είναι απλές και καθαρές, με έμφαση στη λειτουργικότητα. Οι χρήστες θα πρέπει να μπορούν να κατανοούν αμέσως πώς να χρησιμοποιούν την εφαρμογή χωρίς την ανάγκη για εκτενή οδηγίες. Αυτό μπορεί να επιτευχθεί μέσω ενός μινιμαλιστικού σχεδιασμού που εστιάζει στις βασικές λειτουργίες και αποφεύγει την υπερβολική πληροφόρηση (Google, 2021).

**Συνεκτικότητα:** Η συνεκτικότητα σε ολόκληρη την εφαρμογή είναι κρίσιμη για την ευχρηστία. Όλα τα στοιχεία διεπαφής πρέπει να ακολουθούν ένα συνεπές στυλ και να

λειτουργούν με τον ίδιο τρόπο. Αυτό βοηθά τους χρήστες να αισθάνονται άνετα και να κατανοούν γρήγορα πώς να χρησιμοποιούν την εφαρμογή (Tidwell, 2010).

**Ανατροφοδότηση Χρηστών:** Η παροχή ανατροφοδότησης στους χρήστες είναι ουσιαστική για την ευχρηστία. Οι χρήστες πρέπει να γνωρίζουν αμέσως αν οι ενέργειές τους είναι επιτυχημένες ή όχι. Αυτό μπορεί να περιλαμβάνει μηνύματα επιβεβαίωσης, ειδοποιήσεις ή οπτικές ενδείξεις που ενημερώνουν τους χρήστες για την κατάσταση της εφαρμογής (Shneiderman & Plaisant, 2010).

**Διαχείριση Σφαλμάτων:** Μια καλή εφαρμογή πρέπει να μπορεί να διαχειρίζεται τα σφάλματα αποτελεσματικά. Οι χρήστες θα πρέπει να ενημερώνονται για τα σφάλματα με κατανοητό τρόπο και να παρέχονται σαφείς οδηγίες για το πώς να τα διορθώσουν. Αυτό μειώνει την απογοήτευση και αυξάνει την ικανοποίηση των χρηστών (Cooper et al., 2014).

**Προσαρμοστικότητα και Προσβασιμότητα:** Οι εφαρμογές Android πρέπει να είναι προσαρμόσιμες και προσβάσιμες σε όλους τους χρήστες, συμπεριλαμβανομένων των ατόμων με αναπηρίες. Η προσβασιμότητα μπορεί να περιλαμβάνει τη χρήση κειμένου με μεγάλο μέγεθος, υψηλή αντίθεση χρωμάτων και υποστήριξη για τεχνολογίες υποβοήθησης όπως το TalkBack (Google, 2021).

### 4.3 Εργαλεία και Πρακτικές για Βελτίωση της Ευχρηστίας

Η Google παρέχει ένα σύνολο εργαλείων και κατευθυντήριων γραμμών για την ανάπτυξη εύχρηστων εφαρμογών Android. Το Material Design είναι ένα σύνολο κατευθυντήριων γραμμών που παρέχει οδηγίες για τη σχεδίαση οπτικά συνεκτικών και διαισθητικών διεπαφών χρήστη (Google, 2021). Επιπλέον, εργαλεία όπως το Android Studio παρέχουν δυνατότητες για τη δοκιμή και την ανάλυση της ευχρηστίας των εφαρμογών, όπως οι προσομοιώσεις χρήστη και τα εργαλεία ανάλυσης απόδοσης.

Η συνεχής δοκιμή με πραγματικούς χρήστες είναι επίσης κρίσιμη για τη βελτίωση της ευχρηστίας. Οι δοκιμές ευχρηστίας επιτρέπουν στους προγραμματιστές να εντοπίζουν και να διορθώνουν προβλήματα πριν από την κυκλοφορία της εφαρμογής. Οι ανατροφοδοτήσεις



από τους χρήστες μπορούν να παρέχουν πολύτιμες πληροφορίες για το πώς να βελτιωθεί η εμπειρία χρήσης της εφαρμογής (Krug, 2014).

#### **4.4 Σχεδιασμός δοκιμών ευχρηστίας για εφαρμογές Android**

Οι δοκιμές ευχρηστίας εφαρμογών για κινητές συσκευές Android μπορούν να πραγματοποιηθούν με διάφορους τρόπους. Παρακάτω βλέπουμε μερικά βήματα που ακολουθούνται για την πραγματοποίηση δοκιμών ευχρηστίας εφαρμογών Android:

**Ορισμός στόχων:** Καθορισμός στόχων των δοκιμών ευχρηστίας, π.χ. ελέγχου της ευκολίας πλοήγησης, ανταπόκρισης της εφαρμογής, κατανοητότητας των λειτουργιών κ.λπ.

**Σχεδίαση δοκιμασίας:** Δημιουργία ενός σχεδίου για τις δοκιμές ευχρηστίας που θα πραγματοποιηθούν. Καθορισμός των σεναρίων χρήσης που θα δοκιμαστούν, τα βήματα που θα ακολουθηθούν και τα κριτήρια αξιολόγησης.

**Επιλογή εργαλείων:** Επιλογή των κατάλληλων εργαλείων για την πραγματοποίηση των δοκιμών ευχρηστίας. Τέτοια εργαλεία περιλαμβάνουν εργαλεία όπως το UserTesting, το UserZoom, το Lookback και άλλα για την καταγραφή της οθόνης και των αλληλεπιδράσεων των χρηστών (Crowdbotics, 2021).

**Επιλογή δοκιμαστών:** Επιλογή μίας ομάδας δοκιμαστών που θα συμμετάσχουν στις δοκιμές ευχρηστίας. Οι δοκιμαστές θα πρέπει να είναι εκπαιδευμένοι για την εκτέλεση των δοκιμών και να είναι σε θέση να παρέχουν ανταπόκριση και σχολιασμό για την εμπειρία χρήστη (Cruz, Abreu, & Lo, 2019).

**Εκτέλεση των δοκιμών:** Καταγραφή των δραστηριοτήτων των δοκιμαστών καθώς αλληλοεπιδρούν με την εφαρμογή. Παρακολουθήστε τις αντιδράσεις τους, τα σφάλματα που συναντούν και τα σημεία κλειδιά όπου παρουσιάζονται προβλήματα (Harrison, Flood, & Duce, 2013).

**Ανάλυση και αξιολόγηση:** Ανάλυση των δεδομένων που συλλέχθηκαν από τις δοκιμές και αξιολόγηση της απόδοσης της εφαρμογής σε σχέση με τους στόχους που καθορίστηκαν,

καθώς και αναγνώριση των προβλημάτων και τις περιοχές βελτιώσεων (Tarute, Nikou, & Gatautis, 2017).

Οι δοκιμές ευχρηστίας είναι μια σημαντική διαδικασία για τη βελτίωση της εμπειρίας του χρήστη σε μια εφαρμογή Android. Με τη συνεχή ανατροφοδότηση και βελτίωση, είναι δυνατή η δημιουργία μιας εφαρμογής που είναι ευχάριστη και εύχρηστη για τους χρήστες.

#### **4.5 Κλίμακα Likert**

Η κλίμακα Likert, η οποία εισήχθη από τον Rensis Likert το 1932, είναι μια ψυχομετρική κλίμακα που χρησιμοποιείται για τη μέτρηση των στάσεων, απόψεων και αντιλήψεων των ατόμων. Η κλίμακα Likert αποτελείται από μια σειρά δηλώσεων, όπου οι συμμετέχοντες καλούνται να αξιολογήσουν τον βαθμό συμφωνίας ή διαφωνίας τους με κάθε δήλωση σε μια προκαθορισμένη κλίμακα, συνήθως 5 ή 7 βαθμίδων (Likert, 1932).

Ο σκοπός της κλίμακας Likert είναι να παρέχει μια ποσοτική μέτρηση των ποιοτικών δεδομένων, καθιστώντας την ιδανική για έρευνες κοινωνικής επιστήμης, εκπαίδευσης και ψυχολογίας. Η κλίμακα αυτή επιτρέπει στους ερευνητές να αναλύσουν τις τάσεις και τις απόψεις των συμμετεχόντων με μεγαλύτερη ακρίβεια και αξιοπιστία (Bishop & Herron, 2015).

Η ανάλυση δεδομένων που προέρχονται από κλίμακες Likert περιλαμβάνει τη χρήση στατιστικών μεθόδων όπως η μέση τιμή και η ανάλυση διακύμανσης (ANOVA) για τον εντοπισμό διαφορών μεταξύ των ομάδων. Ωστόσο, η χρήση και η ερμηνεία των δεδομένων από κλίμακες Likert πρέπει να γίνεται με προσοχή, καθώς οι διαφορές μεταξύ των βαθμίδων μπορεί να μην είναι ίσες, γεγονός που μπορεί να οδηγήσει σε σφάλματα στην ανάλυση (Sullivan & Artino, 2013).

## 4.6 Ερωτηματολόγιο αξιολόγησης

Παρακάτω βλέπουμε το ερωτηματολόγιο ευχρηστίας που χρησιμοποιήθηκε για την αξιολόγηση ευχρηστίας της εφαρμογής μας.

- 1) Πόσο εύκολη ήταν η διαδικασία εγγραφής ήχου;
- 2) Ήταν εύκολο να ξεκινήσετε και να σταματήσετε μια ηχογράφιση;
- 3) Ήταν ευκολονόητα τα κουμπιά και οι επιλογές για τον έλεγχο της ηχογράφησης;
- 4) Ήταν εύκολο να αναπαραγάγετε μια ηχογράφιση;
- 5) Ήταν ευκολονόητη η διαχείριση των ηχογραφήσεων (π.χ. διαγραφή, μετονομασία, κατηγοριοποίηση);
- 6) Πώς αξιολογείτε την απόκριση της εφαρμογής κατά την εκτέλεση των εντολών;
- 7) Πόσο καλά ενσωματώνεται η εφαρμογή στο γραφικό περιβάλλον του Android;
- 8) Πώς αξιολογείτε την απόδοση της εφαρμογής κατά την εγγραφή μεγάλων διαρκειών;
- 9) Πόσο αξιόπιστη ήταν η εφαρμογή κατά την ηχογράφιση;
- 10) Πώς θα βαθμολογούσατε την συνολική εμπειρία χρήσης της εφαρμογής;

Για τη βαθμονόμηση του ερωτηματολογίου χρησιμοποιήσαμε την κλίμακα Likert με μία πενταβάθμια βαθμονόμηση *καθόλου-λίγο-μέτρια-πολύ-πάρα πολύ*. Για κάθε συμμετέχοντα, πήραμε και κάποια βασικά δημογραφικά στοιχεία όπως, ηλικία, φύλο, μορφωτικό επίπεδο, και γενικό επίπεδο εξοικείωσης με την τεχνολογία. Το ερωτηματολόγιο υλοποιήθηκε και διαμοιράστηκε με τη χρήση google forms και τα δεδομένα συλλέχτηκαν σε ένα google sheet. Το ερωτηματολόγιο βρίσκεται στον παρακάτω σύνδεσμο



Εικόνα 22: QR-code με σύνδεσμο για το ερωτηματολόγιο της εργασίας.

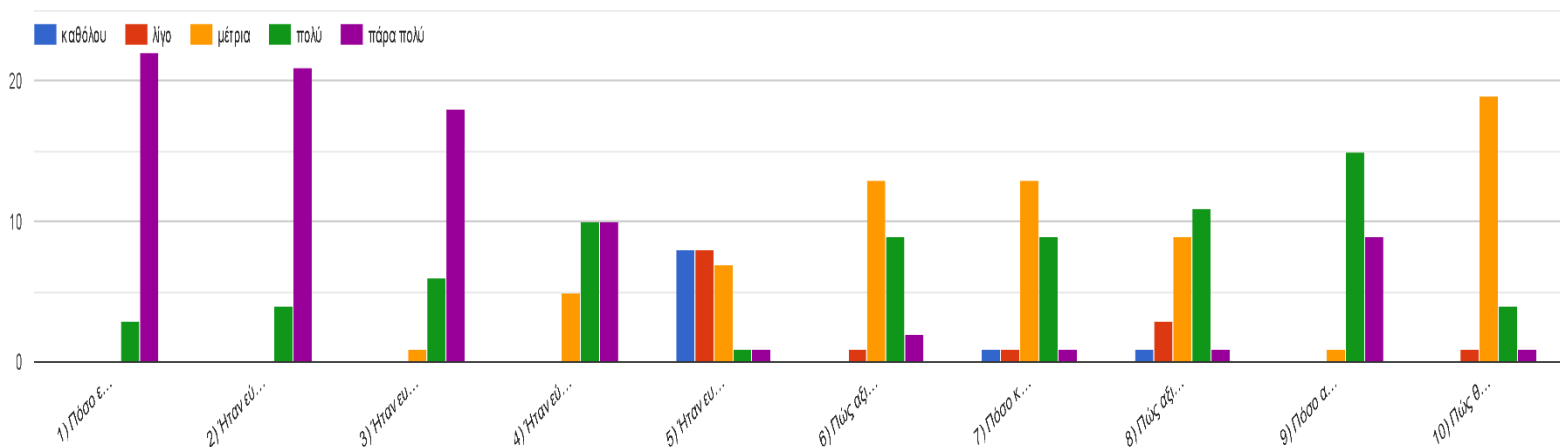
Τέλος, επιπρόσθετα από το ερωτηματολόγιο, ζητήθηκε από 3 expert χρήστες (χρήστες με κάποια επαγγελματική σχέση με τον τομέα είτε του ψηφιακού ήχου είτε της πληροφορικής), να πραγματοποιήσουν τις ίδιες εργασίες με τους άλλους χρήστες αλλά αντί για απλή

συμπλήρωση του ερωτηματολογίου, να μας εκθέσουν τη γνώμη τους σε μία μικρή παράγραφο. Τις απαντήσεις αυτές τις παραθέτουμε στο Παράρτημα Β.

## 4.7 Αποτελέσματα

Το ερωτηματολόγιο μοιράστηκε σε 25 άτομα, στα οποία ζητήθηκε να πραγματοποιήσουν μία ηχογράφιση, να την ακούσουν και να δουν τα αποτελέσματα της επεξεργασίας. Στη συνέχεια τους ζητήθηκε να κάνουν την ίδια εργασία με παλαιότερη ηχογράφιση. Με βάση αυτές τις δύο εργασίες απάντησαν στο ερωτηματολόγιο της προηγούμενης ενότητας και τα αποτελέσματα παρατίθενται στην εικόνα και τον πίνακα που ακολουθεί.

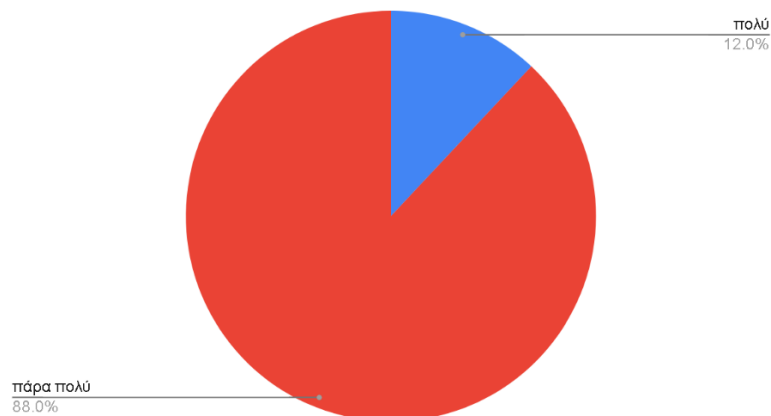
Τα αποτελέσματα συλλέχθηκαν μέσω της διαδικτυακής υπηρεσίας Google Forms υπό τη μορφή γραφήματος και λογιστικού φύλλου. Τα αποτελέσματα για το πρώτο μέρος του ερωτηματολογίου φαίνονται συνολικά στο παρακάτω διάγραμμα.



Εικόνα 23: Αποτελέσματα απαντήσεων ερωτηματολογίου

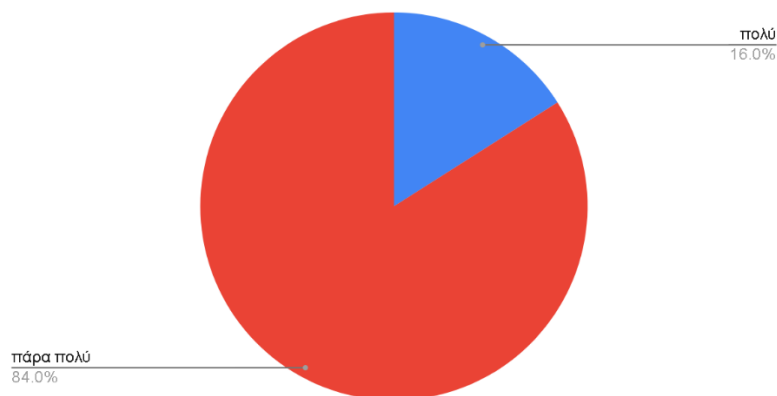
Στα παρακάτω διαγράμματα βλέπουμε την βαθμολόγηση που έκαναν οι χρήστες ανά ερώτηση.

1) Πόσο εύκολη ήταν η διαδικασία εγγραφής ήχου;



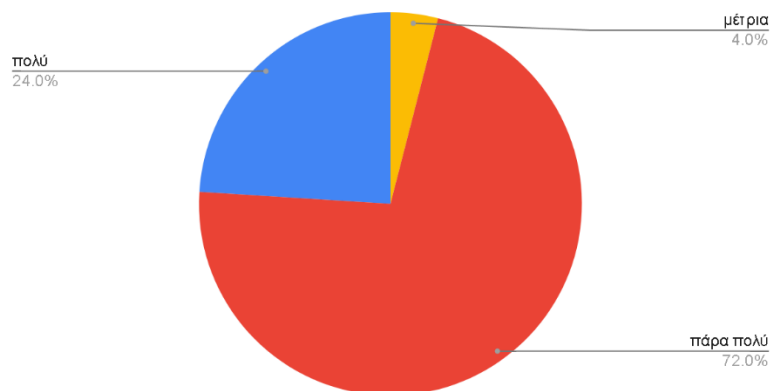
**Εικόνα 24:** διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την πρώτη ερώτηση του ερωτηματολογίου

2) Ήταν εύκολο να ξεκινήσετε και να σταματήσετε μια ηχογράφιση;



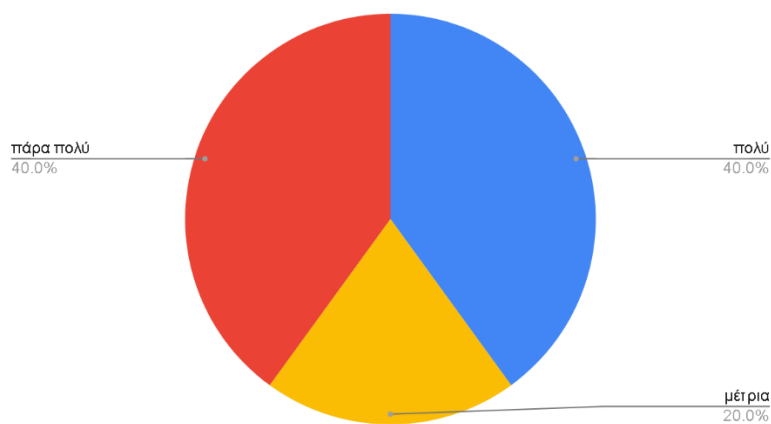
**Εικόνα 25:** διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την δεύτερη ερώτηση του ερωτηματολογίου

3) Ήταν ευκολονόητα τα κουμπιά και οι επιλογές για τον έλεγχο της ηχογράφησης;

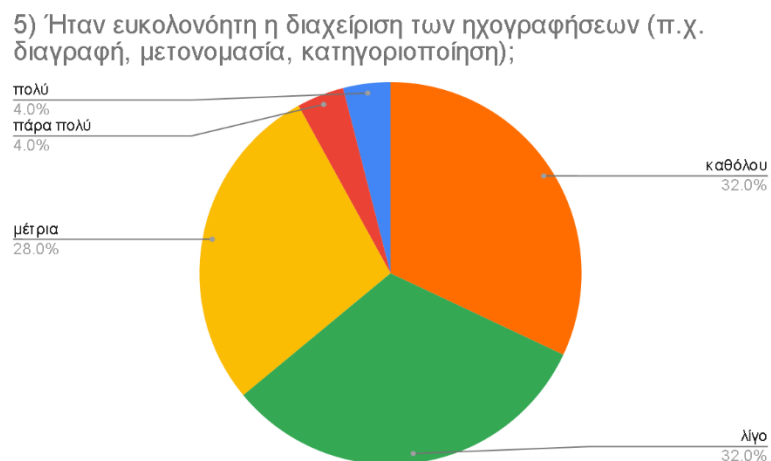


**Εικόνα 26:** διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την τρίτη ερώτηση του ερωτηματολογίου

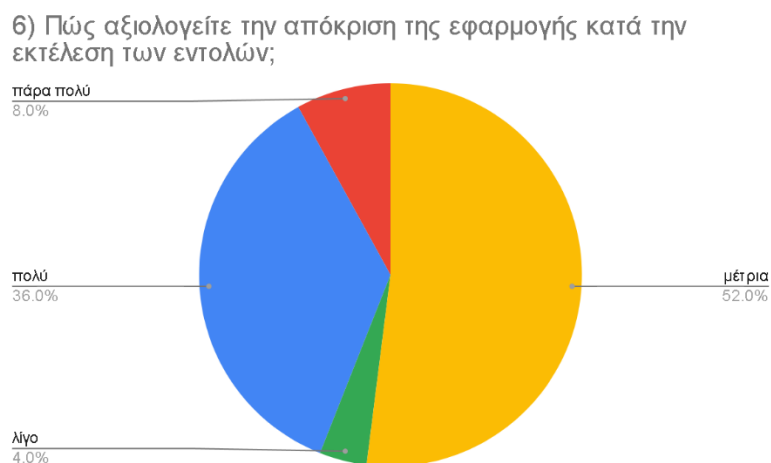
4) Ήταν εύκολο να αναπαραγάγετε μια ηχογράφιση;



**Εικόνα 27:** διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την τέταρτη ερώτηση του ερωτηματολογίου

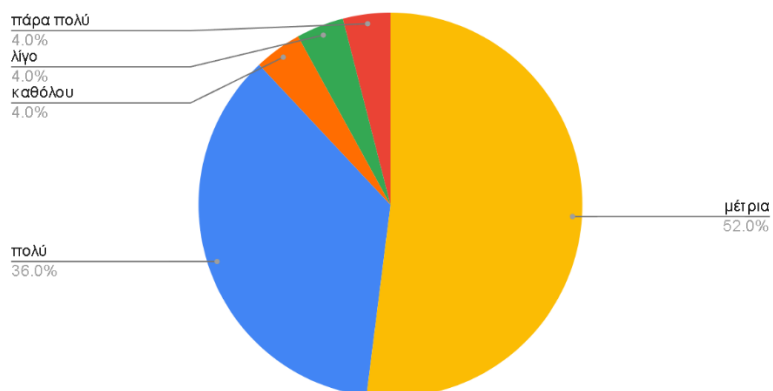


**Εικόνα 28: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την πέμπτη ερώτηση του ερωτηματολογίου**



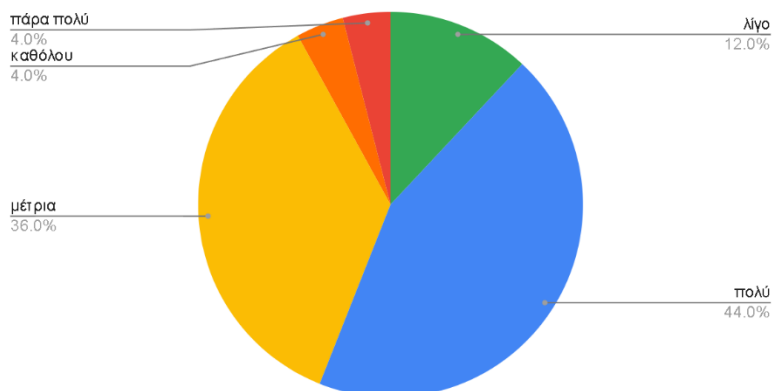
**Εικόνα 29: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την έκτη ερώτηση του ερωτηματολογίου**

7) Πόσο καλά ενσωματώνεται η εφαρμογή στο γραφικό περιβάλλον του Android;



Εικόνα 30:διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την έβδομη ερώτηση του ερωτηματολογίου

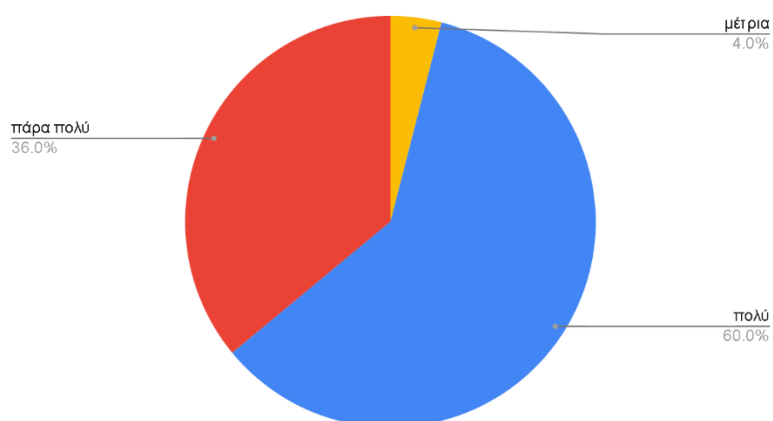
8) Πώς αξιολογείτε την απόδοση της εφαρμογής κατά την εγγραφή μεγάλων διάρκειών;



Εικόνα 31: διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την όγδοη ερώτηση του ερωτηματολογίου

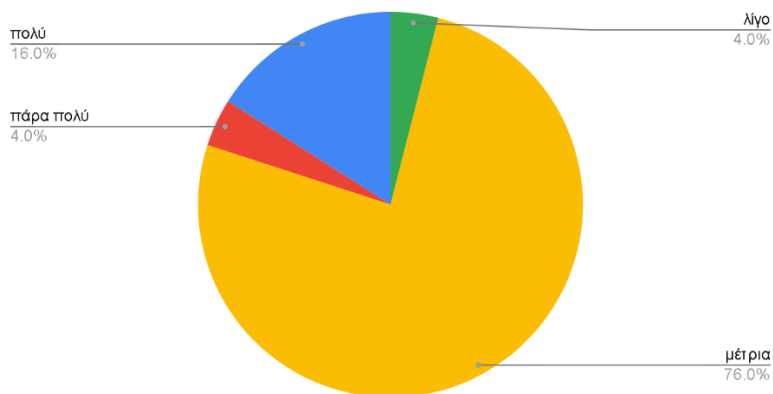


9) Πόσο αξιόπιστη ήταν η εφαρμογή κατά την ηχογράφηση;



**Εικόνα 32:** διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την ένατη ερώτηση του ερωτηματολογίου

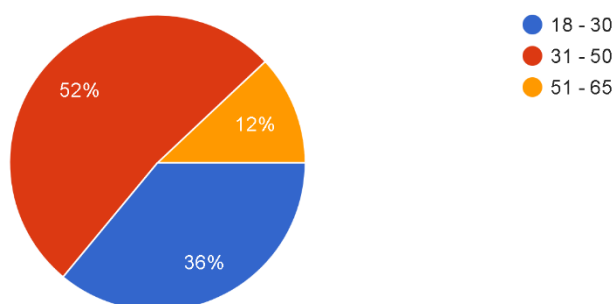
10) Πώς θα βαθμολογούσατε την συνολική εμπειρία χρήσης της εφαρμογής;



**Εικόνα 33:** διάγραμμα με τα αποτελέσματα από τις απαντήσεις των χρηστών για την δέκατη ερώτηση του ερωτηματολογίου

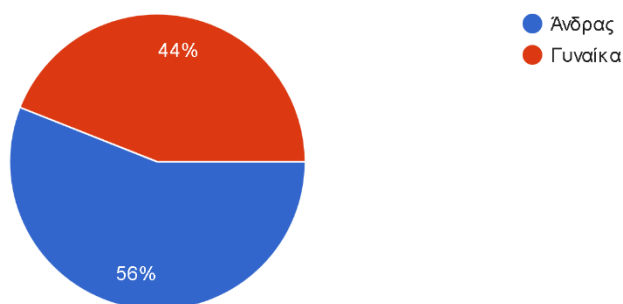
Στα παρακάτω διαγράμματα φαίνονται τα αποτελέσματα από το δεύτερο μέρος του ερωτηματολογίου, όπου συλλέξαμε δημογραφικά στοιχεία για τους συμμετέχοντες.

Ηλικία  
25 responses



**Εικόνα 34:** Διάγραμμα πίτας με την κατανομή των συμμετεχόντων στις τρεις ηλικιακές ομάδες που χρησιμοποιήσαμε.

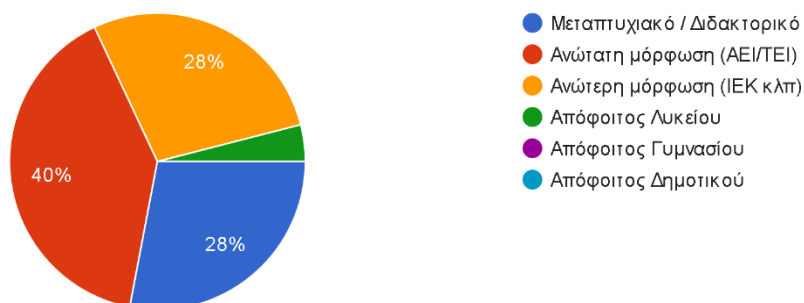
Φύλο  
25 responses



**Εικόνα 35:** Διάγραμμα πίτας με την κατανομή των συμμετεχόντων με βάση τη φύλο τους.

#### Μορφωτικό επίπεδο

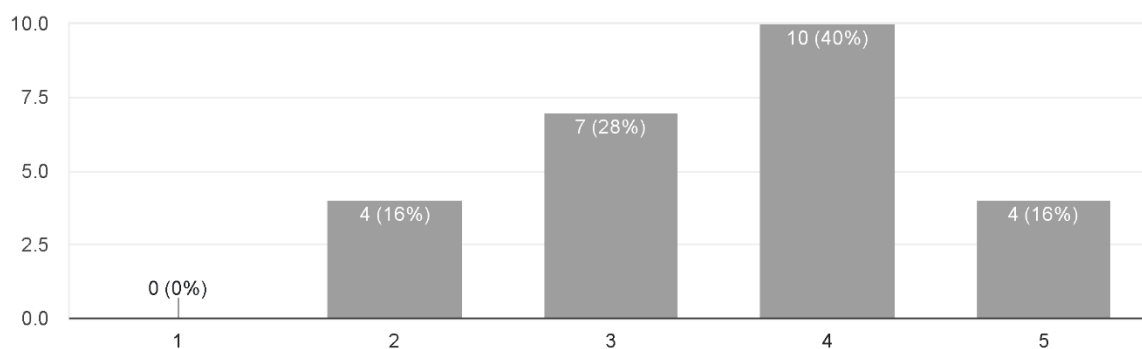
25 responses



Εικόνα 36: Διάγραμμα πίτας με την κατανομή των συμμετεχόντων βάσει του μορφωτικού τους επιπέδου.

#### Άνεση με την τεχνολογία

25 responses



Εικόνα 37: Διάγραμμα, με την κατανομή των συμμετεχόντων βάσει της σχετικής τους άνεσης με την τεχνολογία γενικά.

Τέλος, από τις συνεντεύξεις αποκομίσαμε τέσσερις βασικές παρατηρήσεις για την εφαρμογή. Η πρώτη, αφορά τη διεπαφή και την υπερβολικά απλοϊκή παρουσίασή της. Λειτουργικά όμως, ενώ εκ πρώτης όψεως έδειχνε απλή και ευκολονόητη, παρατηρήθηκε ότι δεν ήταν τελείως ξεκάθαρο τι λειτουργία κάνει κάθε κουμπί σε κάθε σκέλος της λειτουργικότητας. Δεύτερη παρατήρηση που έγινε ήταν ότι φαινόταν να παγώνει η εφαρμογή κατά τη διάρκεια της επεξεργασίας της ηχογράφησης, γεγονός που δημιουργούσε ανασφάλεια στους χρήστες. Δεν υποδεικνυόταν δηλαδή μέσω της διεπαφής ότι κάτι βρισκόταν σε εξέλιξη, πέρα από το κουμπί του process που φαινόταν κολλημένο στην πατημένη θέση. Η τρίτη παρατήρηση, αφορούσε την απουσία κάποιου player για την αναπαραγωγή της ηχογράφησης με δυνατότητες έναρξης, παύσης, διακοπής, και ένδειξη του σημείου εντός της συνολικής διάρκειας που βρίσκεται η αναπαραγωγή (seekbar). Τέλος, η τέταρτη παρατήρηση αφορούσε την λειτουργία του κουμπιού load το οποίο είχε πολλαπλά προβλήματα λόγω της ελλιπούς υλοποίησης με αποτέλεσμα αρκετές από τις φορές που χρησιμοποιούνταν να προκαλεί το κλείσιμο της εφαρμογής.

## 4.8 Συζήτηση

Συνολικά, εκ πρώτης όψεως, παρατηρούμε ότι στην πρώτη εργασία που είχαν οι χρήστες, όλα ήταν εύκολα όπως φαίνεται και από τις απαντήσεις τους, ενώ η αποτελεσματικότητα της εφαρμογής μειώθηκε δραματικά στην δεύτερη εργασία. Κάτι τέτοιο ήταν αναμενόμενο σε κάποιο βαθμό, γιατί βασική προτεραιότητα της εφαρμογής ήταν να δείξει ένα τρόπο με τον οποίο, κάποιος κάνει μια ηχογράφηση, την ακούει και τέλος παράγει μιας μορφής τεχνική αξιολόγηση της ποιότητας της ηχογράφησης με αυτόματο τρόπο.

Σε ένα ολοκληρωμένο προϊόν, όπως είναι το jReporter (Vrysis et al., 2019), καθώς και πληθώρα άλλων εφαρμογών ηχογράφησης, προσφέρουν δυνατότητες διαχείρισης και αρχειοθέτησης ηχογραφήσεων, κάτι που υλοποιήθηκε σε κάποιο βαθμό, αλλά δεν αποτέλεσε προτεραιότητα σε αυτή την εργασία.

Ως γραφική διεπαφή, η εφαρμογή μας αξιοποίησε μόνο υπάρχοντα στοιχεία που προσέφερε η ίδια η πλατφόρμα του Android αφού και πάλι δεν ήταν τόσο ζητούμενο της εργασίας η

διεπαφή να έχει κάποια συγκεκριμένη μορφή, αλλά περισσότερο να παρέχει με τον απλούστερο δυνατό τρόπο μια λειτουργικότητα.

Τα δημογραφικά στοιχεία που συλλέξαμε, μας δίνουν και αυτά μια εικόνα για την αποτελεσματικότητα της εφαρμογής μας, με βάση όμως τους χρήστες που τη χρησιμοποίησαν. Παρακάτω θα εκθέσουμε κάποιες παρατηρήσεις που κάναμε για το δείγμα μας, με βάση την κάθε παράμετρο που μετρήσαμε.

Με βάση την ηλικία, παρατηρούμε ότι η ηλικιακή ομάδα 31-50 είχε την μεγαλύτερη άνεση με την τεχνολογία, αλλά όχι με μεγάλη διαφορά από τις άλλες δύο ομάδες. Τα δείγματα που έχουν το μεγαλύτερο βαθμό, περιλαμβάνουν τους τρεις ειδικούς, στους οποίους ζητήσαμε να μας πουν και την γνώμη τους για την εφαρμογή. Την καλύτερη βαθμολογία την έδωσε η ίδια ηλικιακή ομάδα, ενώ τη χειρότερη η 18 - 30. Σημειώνουμε ότι το δείγμα μας, εκτός από μικρό μέγεθος, γενικότερα, δεν περιείχε μεγάλη ποικιλομορφία, και ειδικά για τις ηλικίες, η ομάδα 51 – 65 είχε μόνο τρία άτομα, με αποτέλεσμα να μην μπορούμε να εξάγουμε ιδιαίτερα αξιόπιστα συμπεράσματα βάσει της ηλικίας των συμμετεχόντων.

Με βάση το φύλο, παρατηρούμε ότι οι γυναίκες του δείγματός μας είχαν μικρότερη άνεση με την τεχνολογία σε σχέση με τους άνδρες, και έδωσαν συνολικά χαμηλότερη βαθμολογία. Τέλος, όλοι έδειξαν ότι δυσκολεύτηκαν στο κομμάτι της φόρτωσης αρχείου και της διαχείρισης των ηχογραφήσεων.

Με βάση το μορφωτικό επίπεδο, τα άτομα του δείγματός μας τουλάχιστον ανώτερη μόρφωση (μεταδευτεροβάθμια, ΙΕΚ, κλπ.), με εξαίρεση έναν απόφοιτο λυκείου, με την πλειοψηφία να έχει ανώτατη μόρφωση (απόφοιτοι ΑΕΙ/ΤΕΙ). Σε αυτή την παράμετρο το δείγμα μας είχε μια κάποια ποικιλομορφία, ενώ επιπλέον κάποια από τα άτομα είχαν τίτλους εκτός του τεχνολογικού τομέα. Τα άτομα με την μεταπτυχιακό/διδακτορικό, βαθμολόγησαν καλύτερα από όλους την εφαρμογή μας, ενώ οι απόφοιτοι ανώτερης εκπαίδευσης, έδωσαν τη χαμηλότερη βαθμολογία, παρόλο που είχαν παρόμοιο επίπεδο εξοικείωσης με την τεχνολογία. Ο απόφοιτος λυκείου που έδωσε την χαμηλότερη συνολική βαθμολογία είχε χαμηλή άνεση με την τεχνολογία, αλλά όντας ένα μόνο άτομο, δεν αποτελεί αξιόπιστη πηγή για συμπεράσματα.

Για την παράμετρο της εξοικείωσης/άνεσης με την τεχνολογία, παρατηρούμε ότι οι περισσότεροι είχαν αρκετά καλή εξοικείωση με την τεχνολογία (4 στα 5 της κλίμακας μας), ενώ επίσης φαίνεται ότι αυτοί με την καλύτερη εξοικείωση, έδωσαν και τη μεγαλύτερη βαθμολογία. Φαινόταν δηλαδή από τα δεδομένα ότι η βαθμολογία που δόθηκε ήταν ανάλογη της άνεσης με την τεχνολογία.

Τέλος, στο ερώτημα για τη συνολική εμπειρία από το πρώτο μέρος, παρατηρούμε ότι οι περισσότεροι, το 76%, του δείγματος είχαν μέτρια εμπειρία. Το ποσοστό αυτό επίσης περιέχει και το σύνολο της ποικιλομορφίας του δείγματός μας σε όλες τις δημογραφικές παραμέτρους, πράγμα που του δίνει περισσότερο βάρος σαν δεδομένο. Στο ίδιο ερώτημα επίσης, την καλύτερη βαθμολογία την έδωσαν οι συμμετέχοντες με τη μεγαλύτερη εξοικείωση με την τεχνολογία.

Αξίζει να σημειωθεί ότι οι περισσότεροι, 31-50 άνδρες του δείγματός μας, ήταν τεχνικοί του χώρου της οπτικοακουστικής κάλυψης εκδηλώσεων (ηχολήπτες, φωτιστές, τεχνικοί εικόνας και βίντεο, κλπ.), γεγονός που απομακρύνει περισσότερο το δείγμα μας από την εικόνα του γενικού πληθυσμού ή του κοινού των δημοσιογράφων που στόχευαν σχετικές εργασίες, όπως των (Vrysis et al., 2019).

## 5 Συμπεράσματα και μελλοντικές κατευθύνσεις

Συνολικά οι ηχογραφήσεις είχαν βαθμολογία από τον μηχανισμό της εφαρμογής 1 και 2, δηλαδή ιδιαίτερα θορυβώδεις. Αυτό συνέβη για δύο λόγους, πρώτον επειδή δεν υπάρχει δυνατότητα απενεργοποίησης του AGC (Automatic Gain Controller) για το μικρόφωνο της συσκευής, οπότε με το να μεγιστοποιείται το κέρδος σήματος στο μικρόφωνο για μέγιστη απολαβή, έχουμε και μέγιστη απολαβή περιβαλλοντικού θορύβου (θορύβου βάθους δωματίου κλπ.). Για το λόγο αυτό, δεν πήραμε ποτέ τιμές SFM κάτω από 0.80, ενώ ταυτόχρονα, ο συνδυασμός των χαρακτηριστικών της κάψας του μικροφώνου με μεγάλη απολαβή υψηλομεσσαίων και υψηλών συχνοτήτων, οι τιμές του τονικού κέντρου (spectral centroid) του σήματος ήταν πάντα αρκετά υψηλότερα από το αναμενόμενο, και ιδιαίτερα όταν στην ομιλία υπήρχαν σύμφωνα και φωνήματα όπως «σ», «τσ», «ψ». Ενώ τέλος, στον έλεγχο για ψαλιδισμό, δεν μπορέσαμε ποτέ να προκαλέσουμε το φαινόμενο αυτό, ξανά, λόγω του AGC και του όποιου άλλου αυτοματισμού ελέγχου δυναμικού εύρους (auto-limiting) το οποίο, όπως παρατηρήσαμε σε δοκιμές, έμπαινε σε λειτουργία τόσο γρήγορα που δεν παρατηρήσαμε ψαλιδισμό στα δεδομένα ακόμα και όταν χτυπούσαμε το τηλέφωνο στο σημείο του μικροφώνου ή ακόμα και όταν το εκθέταμε σε πολύ υψηλής στάθμης ήχους. Για την επίδειξη του ελέγχου αυτού όμως, θέσαμε ένα κατώφλι αρκετά χαμηλά, στα -3.5 dBFS, ώστε όταν έχουμε ήχους με κορυφές σε αυτή τη στάθμη, που είναι επαρκώς υψηλή, να έχουμε ένδειξη ψαλιδισμού και στάθμης που τον προκάλεσε.

Ως μελλοντικές κατευθύνσεις μπορούν να τεθούν, κατ' αρχήν από τις παρατηρήσεις των ειδικών χρηστών και τα αποτελέσματα του ερωτηματολογίου: η υλοποίηση μιας πιο κατανοητής και ολοκληρωμένης διεπαφής χρήστη, η δυνατότητα παρουσίασης, διαχείρισης, και ακρόασης προγενέστερων ηχογραφήσεων, η υλοποίηση ενός ολοκληρωμένου μηχανισμού αναπαραγωγής των ηχογραφήσεων (media player) με όλες τις βασικές λειτουργίες του, η δυνατότητα εγγραφής, επεξεργασίας και αποθήκευσης σε codecs συμπιεσμένου ψηφιακού ήχου, και η δυνατότητα σύνδεσης εξωτερικού ενσύρματου ή ασύρματου μικροφώνου. Επιπλέον η παρουσίαση των αποτελεσμάτων μπορεί μελλοντικά, με βελτιώσεις/προσθήκες, να περιέχει προτάσεις για τη βελτίωση της όποιας εγγραφής αλλά και επιλογή για την αυτόματη διόρθωση των όποιων ατελειών ανιχνεύθηκαν όπως είδαμε σε

σχετικές εργασίες στο δεύτερο κεφάλαιο. Τέλος, ο αλγόριθμος της εφαρμογής στα διάφορα σκέλη της μπορεί να βελτιωθεί ώστε να συμπεριλάβει περισσότερα φασματικά χαρακτηριστικά, για τη βελτίωση της ανίχνευσης προβλημάτων, αλλά και τον αυτοματισμό εργασιών όπως κατηγοριοποίηση και η επισύναψη ετικετών στα αρχεία που καταγράφονται με σκοπό την καλύτερη οργάνωση και διαχείρισή τους.



## Βιβλιογραφία

Agostini, G., Longari, M., & Pollastri, E. (2003). Musical Instrument Timbres Classification with Spectral Features. EURASIP Journal on Advances in Signal Processing, 2003(943279). <https://doi.org/10.1155/S1110865703210118>

All About Circuits. (2021). Understanding Bandwidth in Communications and Computing. Ανακτήθηκε από All About Circuits(<https://www.allaboutcircuits.com/technical-articles/understanding-bandwidth-in-communications-and-computing/>).

Aquila DSP. (2023). Aquila - a C++ library for digital signal processing. Retrieved from Aquila DSP

*Bandwidth of Signals: Definition, Concepts, Applications, Videos, Examples.* (n.d.). Retrieved September 17, 2024, from <https://www.toppr.com/guides/physics/communication-systems/bandwidth-of-signals/>

Bishop, P. A., & Herron, R. L. (2015). Use and misuse of the Likert item responses and other ordinal measures. International Journal of Exercise Science, 8(3), 297-302.

Βρύζας, Ν., Σιδηρόπουλος, Ε., Βρύσης, Λ., Αβραάμ, Ε., & Δημουλάς, Χ. (2019). Machine-assisted reporting in the era of mobile journalism: the MOJO-mate platform.

Βρύσης, Λ. Δ. (2014). *Ανάπτυξη περιβάλλοντος σημασιολογικής διαχείρισης ηχητικής πληροφορίας* (Doctoral dissertation, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης).

Chen, J., Huang, Y. A., Li, Q., & Paliwal, K. K. (2004). Recognition of noisy speech using dynamic spectral subband centroids. *IEEE Signal Processing Letters*, 11(2 PART II), 258–261. <https://doi.org/10.1109/LSP.2003.821689>

Crowdbotics. (2021). A Guide to Mobile App Usability Testing: Methods, Tools, and Best Practices. Ανακτήθηκε από <https://crowdbotics.com/posts/blog/a-guide-to-mobile-app-usability-testing-methods-tools-and-best-practices/>

Cruz, L., Abreu, R., & Lo, D. (2019). To the attention of mobile software developers. *Empirical Software Engineering*, 24(4), 2438–2468. <https://doi.org/10.1007/S10664-019-09701-0>

*Designing User Interface With Views In Android Applications*. (n.d.). Retrieved September 14, 2024, from <https://www.c-sharpcorner.com/article/designing-user-interface-with-views-in-android-application/>

*Develop a UI with Views | Android Studio | Android Developers*. (n.d.). Retrieved September 14, 2024, from <https://developer.android.com/studio/write/layout-editor>

GitHub. (2023). JTransforms. Retrieved from <https://github.com/wendykierp/JTransforms>

GTE Media. (2023). The Evolution of Portable Media Players: From the Walkman to the Digital Era. Retrieved from GTE Media

Guillemard, M., Ruwwe, C., & Zölzer, U. (n.d.). *j-DAFx - Digital Audio Effects in Java*. <http://ant.hsu-hh.de>

Hansen, J. H. L., Stauffer, A., & Xia, W. (2018). Nonlinear waveform distortion: Assessment and detection of clipping on speech data and systems. *Journal of the Acoustical Society of America*, 144(3\_Supplement), 1871. <https://doi.org/10.1121/1.5068230>

Harrison, R., Flood, D., & Duce, D. (2013). *Usability of mobile applications: literature review and rationale for a new usability model*.  
<http://www.journalofinteractionscience.com/content/1/1/1>

*History of Android - GeeksforGeeks*. (n.d.). Retrieved September 14, 2024, from  
<https://www.geeksforgeeks.org/history-of-android/>

JorenSix/TarsosDSP. (2023). TarsosDSP, a Real-Time Audio Processing Framework in Java. Retrieved from <https://github.com/JorenSix/TarsosDSP>

*Music Recognition Algorithms: How Does Shazam Work? | Toptal®*. (n.d.). Retrieved September 14, 2024, from <https://www.toptal.com/algorithms/shazam-it-music-processing-fingerprinting-and-recognition>

Laguna, C., & Lerch, A. (2016). *An Efficient Algorithm for Clipping Detection and Declipping Audio*. <http://www.aes.org/e-lib>.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 55.

Müller, M. (2015). Fundamentals of Music Processing. *Fundamentals of Music Processing*.  
<https://doi.org/10.1007/978-3-319-21945-5>

Peeters, G. (2004). *A large set of audio features for sound description (similarity and classification) in the CUIDADO project*.

Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., & McAdams, S. (2011). The Timbre Toolbox: Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America*, 130(5), 2902–2916. <https://doi.org/10.1121/1.3642604>

Santucci, R. & Gupta, T. & Shah, Mohit & Spanias, Andreas. (2010). Advanced functions of Java-DSP for use in electrical and computer engineering senior level courses.

Sarti, Augusto, Udo Zoelzer, Xavier Serra, Mark Sandler, and Simon Godsill. 2011. “Digital Audio Effects.” *EURASIP Journal on Advances in Signal Processing* 2010 2010:1 2010(1):1–2. doi: 10.1155/2010/459654.

Anon. n.d. “Digital Signal Processing: Sampling Rates, Bandwidth, Spectral Lines, and More....” Retrieved September 14, 2024 (<https://community.sw.siemens.com/s/article/digital-signal-processing-sampling-rates-bandwidth-spectral-lines-and-more>).

Six, Joren. 2015. *Digital Sound Processing and Java Documentation for the TarsosDSP \* Audio Processing Library*.

Spanias, A., Painter, T., & Atti, V. (1963). *AUDIO SIGNAL PROCESSING AND CODING*.

Sullivan, G. M., & Artino, A. R., Jr. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education*, 5(4), 541-542.

Tarute, A., Nikou, S., & Gatautis, R. (2017). Mobile application driven consumer engagement. *Telematics and Informatics*, 34(2), 145–156.

Test and Measurement Tips. (2021). \*Bandwidth basics and fundamentals\*. Ανακτήθηκε από  
Test and Measurement Tips(<https://www.testandmeasurementtips.com/bandwidth-basics-and-fundamentals/>).

The Saturday Evening Post. (2023). From Walkman to Your Phone: 40 Years of Portable  
Music. Retrieved from The Saturday Evening Post  
(<https://www.saturdayeveningpost.com/2020/06/from-walkman-to-your-phone-40-years-of-portable-music/>)

Unison. (2021). What Is Audio Clipping? Enhance Your Music with Pristine Clarity.  
Retrieved from Unison Audio

Vrysis, Lazaros & Vryzas, Nikolaos & Sidiropoulos, Efstathios & Avraam, Evangelia &  
Dimoulas, Charalampos. (2019). jReporter: A Smart Voice-Recording Mobile Application.

Zölzer, Udo., & Amatriain, Xavier. (2002). *DAFX: digital audio effects*. Wiley.

## Παράρτημα Α: Σύνδεσμος με την πλήρη λίστα κώδικα της εφαρμογής

Η πλήρης λίστα του κώδικα της εφαρμογής φιλοξενείται στον ιστότοπο github.com, στη διεύθυνση:

<https://github.com/giannislel/AndroidAudioRecordingEvaluator.git>

Στο Android Studio πηγαίνετε στο μενού File -> New -> Project From Version Control

Εκεί εισάγετε τον παραπάνω σύνδεσμο, και φορτώνεται ένα αντίγραφο του project στον υπολογιστή σας.

## Παράρτημα Β: Απαντήσεις εξειδικευμένων χρηστών

### Χρήστης Α

- Το πλήκτρο "Stop" έπρεπε να είναι μαζί με το "Record" ή το "Record" κατά τη διάρκεια της εγγραφής να γίνεται "Stop". Το ίδιο ισχύει και για το ζεύγος "Play" - "Stop".
- Τα κουμπιά ας είχαν περιθώρια αριστερά και δεξιά.
- Κατά τη διάρκεια του "processing" μερικές φορές η διαδικασία παίρνει χρόνο, οπότε ένα popup παράθυρο που να λέει "processing" θα βοηθούσε.
- Στον χώρο κάτω από τα κουμπιά, στις ενδείξεις, καλό θα ήταν οι ενδείξεις να μην είναι σε checkboxes, αλλά κάπως αλλιώς, γιατί έτσι νομίζει ο χρήστης ότι "πατιούνται" και αργεί να καταλάβει ότι είναι το αποτέλεσμα της ανάλυσης.
- Θα ήταν βοηθητικό στο label που δείχνει το path του αρχείου, να ήταν κάπως διαφορετικό το GUI.

### Χρήστης Β

- Το συνολικό design είναι πολύ βασικό
- Τα βήματα για τη λειτουργία της εφαρμογής μπορεί να μην είναι και τόσο σαφή για κάποιους χρήστες.
- Δείχνει να παγώνει όταν πατιέται το process. Καλό θα ήταν να υπάρχει κάποια ενημέρωση όταν πατιέται το πλήκτρο, τύπου «η επεξεργασία ενδέχεται να διαρκέσει κάποια λεπτά»
- Κατά τη διάρκεια της εγγραφής δεν υπάρχει κάποια ένδειξη ότι συμβαίνει κάτι, πχ ένα χρονόμετρο που δείχνει την τρέχουσα διάρκεια της ηχογράφησης.
- Το playback control είναι πολύ βασικό, θα βοηθούσε πολύ ένας απλός player με τα βασικά controls.
- Θα βοηθούσε κάποιο κουμπί που να ανοίγει τον φάκελο που περιέχει το αρχείο που γράφτηκε, αντί να το ψάχνει ο χρήστης.

## Χρήστης Γ

- Ξεκινώντας από το Load, η εφαρμογή κόλλησε και έκλεισε στη δεύτερη προσπάθεια φόρτωσης του ίδιου αρχείου. Προσπαθώντας εκ νέου μετά από ηχογράφιση, κλείσιμο και ξανά άνοιγμα της εφαρμογής, δείχνει να δουλεύει κανονικά.
- Δεν δείχνει με κάποιο τρόπο ότι συμβαίνει κάτι είτε όταν πατιέται το record, είτε όταν πατιέται το play.
- Το playback έχει πρόβλημα, αν δεν πατηθεί το stop μετά το play δεν μπορεί να ξαναγίνει αναπαραγωγή του ίδιου, ήδη φορτωμένου, αρχείου.
- Θα ήταν καλό να υπήρχε πιο εύκολη πρόσβαση στο αρχείο της ηχογράφησης.



## Παράρτημα Γ: Σύνδεσμος με πίνακα αποτελεσμάτων ερωτηματολογίου



Εικόνα 38: Σύνδεσμος με πίνακα αποτελεσμάτων ερωτηματολογίου

Υπέθυνη Δήλωση Συγγραφέα:

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1986 και τα άρθρα 2,4,6 παρ. 3 του Ν. 1256/1982, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής εργασίας και δεν προσβάλλει κάθε μορφής πνευματικά δικαιώματα τρίτων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον.