



Ψηφιακές Αναπαραστάσεις της Μουσικής

Διάλεξη 1: Εισαγωγή

Χ. Αλεξανδράκη

Αντικείμενο Μαθήματος

❖ Το μάθημα «Ψηφιακές Αναπαραστάσεις της Μουσικής» ασχολείται με κάθε μορφής περιγραφική πληροφορία που μπορεί να συσχετιστεί με τη μουσική.

- Τέτοια πληροφορία είναι για παράδειγμα, τα **μεταδεδομένα (metadata)** που περιγράφουν ένα μουσικό έργο συνολικά, καθώς και η **περιγραφή του περιεχομένου του (music content description)**, αναφορικά με τις νότες, τις ρυθμικές αξίες και τα μουσικά όργανα.

Φόρτος εργασίας και αξιολόγηση/ βαθμολογία

❖ Παρακολούθηση και εβδομαδιαίος φόρτος εργασίας

- 2 ώρες Θεωρίας
- 2 ώρες Άσκηση
- 3-4 ώρες μελέτη στο σπίτι

❖ Η επίδοσή σας στο μάθημα θα αξιολογηθεί με δύο τρόπους

- 1. Εξετάσεις προόδου (40%):** Θα πραγματοποιηθούν 2-3 εξετάσεις προόδου κατά τη διάρκεια του εξαμήνου στο Εργαστήριο Ηλεκτρονικών Υπολογιστών
- 2. Τελική γραπτή εξέταση θεωρίας (60%):** Η τελική εξέταση θα πραγματοποιηθεί στο τέλος του εξαμήνου και θα είναι γραπτή εξέταση, στο χώρο του Αμφιθεάτρου

❖ Επαναξιολόγηση

- Στην εξεταστική του Σεπτεμβρίου, μπορείτε να αξιολογηθείτε **μόνο για τη γραπτή εξέταση**, δηλαδή μόνο το 60% του τελικού σας βαθμού

Συναφείς Επιστημονικές Κατευθύνσεις

- ❖ Σήμερα, οι αναπαραστάσεις αυτές αποτελούν κεντρικό αντικείμενο έρευνας, τόσο σε μεγάλους τεχνολογικούς κολοσσούς (Google, Apple κ.λπ.) όσο και σε επιστημονικές περιοχές όπως:
 - Υπολογιστική Μουσικολογία (Computational Musicology)
 - Η ανάλυση της μουσικής με υπολογιστικές μεθόδους
 - Ανάκτηση Μουσικής Πληροφορίας (Music Information Retrieval)
 - Η αυτόματη εξαγωγή σημασιολογικής πληροφορίας από μουσικά αρχεία
 - Τεχνητή νοημοσύνη στη Μουσική
 - Η ανάπτυξη αυτόνομων εφαρμογών που διαθέτουν μουσική παιδεία, π.χ. μπορούν να εκτελέσουν, να αναλύσουν, να συνθέσουν και να επεξεργαστούν μουσικά έργα

Μουσική Πληροφορία

❖ Οποιαδήποτε πληροφορία σχετίζεται με κάποιο μουσικό έργο (εκτός του ηχητικού αρχείου)

❖ Δύο είδη:

1. Μακροσκοπικά Μεταδεδομένα

- Μία τιμή για όλο το μουσικό κομμάτι, της μορφής
 - *key = value*

2. Περιγραφή Μουσικού περιεχομένου

- Τα δομικά στοιχεία του έργου αναφορικά με νότες συγχορδίες, ρυθμό, δυναμικές και μουσικά όργανα



Μακροσκοπικά Μεταδεδομένα

Μακροσκοπικά Μεταδεδομένα

❖ Γενικά:

➤ *Metadata* ≡ *Data About Data*

❖ Στη μουσική:

➤ Αναγνωριστική ή/και περιγραφική πληροφορία ενός μουσικού έργου

Πρώτη μορφή μουσικών μεταδεδομένων

❖ Πρωτοεμφανίζονται ως το CD-Text

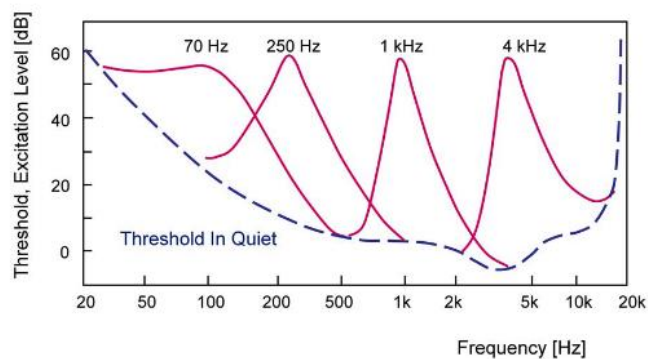
- Προστέθηκε στο πρωτόκολλο Red Book Compact Disk το 1996
- Δεδομένα κειμένου που αναφέρονται είτε συνολικά στο album, είτε σε κάθε ξεχωριστό CD Track:
 - track name, track number, album, artist, year, κ.ο.κ.

Μια μικρή ιστορία για τα μεταδεδομένα (1/2)

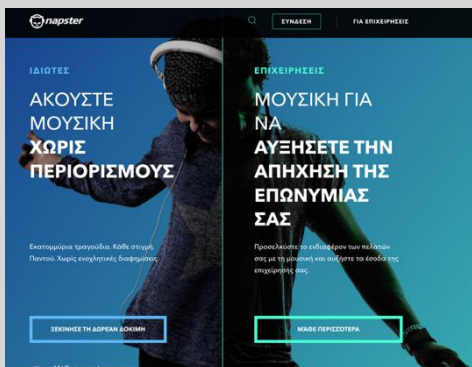
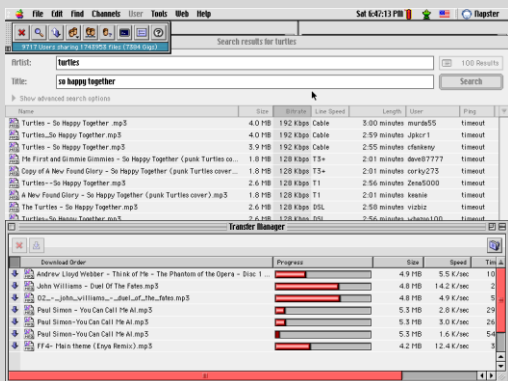


- 1995 - Δημοσίευση Προτύπου MP3
 - **MPEG-1 Audio Layer III**
 - Fraunhofer Institute of Technology (με τη σημερινή επωνυμία **Fraunhofer Society**)
- Απωλεστική συμπίεση (**Lossy Compression**)
 - αξιοποιεί ψυχοακουστικά φαινόμενα (**auditory masking**)
 - επιτυγχάνει πολύ μεγάλους λόγους συμπίεσης, της τάξης 11:1

Critical Bands



Μια μικρή ιστορία για τα μεταδεδομένα (2/2)



- 1999 – λειτουργία του Napster
 - Peer-to-peer δίκτυο για ανταλλαγή αρχείων MP3
 - Αναπτύχθηκε από τους Shawn Fanning και Sean Parker
 - Έφτασε τους 80 εκατομμύρια χρήστες!!
- 2000 – Metallica
 - πρώτη αγωγή στο napster μέσω του οποίου κυκλοφορούσε demo πριν την επίσημη κυκλοφορία του
 - Αγωγή από διάφορες δισκογραφικές για παραβίαση πνευματικών δικαιωμάτων
- 2001 – κλείσιμο του Napster
 - Λόγω αδυναμίας να συμμορφωθεί με ζητήματα πνευματικής ιδιοκτησίας
- 2002 –εξαγοράσθηκε από την εταιρία Roxio
 - Ως ηλεκτρονικό κατάστημα μουσικής (doenload/streaming)
- Σήμερα
 - Πλατφόρμα για συνεργασία μέσω μουσικών με αξιοποίηση τεχνολογιών με AI και immersion

Μεταδεδομένα

❖ Το mp3 και το Napster:

- Επαναπροσδιόρισαν τον τρόπο διανομής μουσικού περιεχομένου
 - Αρχικά με πειρατική ανταλλαγή αρχείων και εν συνεχεία με νόμιμες υπηρεσίες
- Πλέον τα μουσικά αρχεία ήταν διάσπαρτα στο Διαδίκτυο
- **Πρόβλημα:**
 - Απουσία πληροφορίας που περιγράφει τα αρχεία!!
 - Λόγω πειρατικής ανταλλαγής, το όνομα αρχείου (filename) μπορεί να ήταν λανθασμένο....

Λύση: το πρότυπο ID3

❖ 1996 - Συμπλήρωση MP3 αρχείων με:

➤ Πεδίο μεγέθους 128 Bytes που περιλαμβάνει βασικές **ετικέτες (tags)**:

- Artist
- Album
- Year
- Genre
- Comment

Εξέλιξη: ID3v2

❖ Οι ετικέτες

- προσαρτώνται στην αρχή του αρχείου MP3 και όχι στο τέλος, ώστε να εμφανίζονται σε media players χωρίς να απαιτείται η πλήρης φόρτωση του αρχείου (π.χ. Audio streaming)
- Είναι μεταβλητού μεγέθους με ανώτατο όριο 256MB
- Περιλαμβάνουν πρόσθετες πληροφορίες, όπως:
 - Εξώφυλλο του άλμπουμ (Cover art)
 - Στίχους (Lyrics)
 - Στοιχεία για πνευματικά δικαιώματα (Copyrights)



Μακροσκοπικά Μεταδεδομένα σε Ετικέτες

- ❖ Σήμερα χρησιμοποιούνται και σε άλλα File Formats (πέραν του MP3), π.χ.
 - aiff, wav
- ❖ Ορισμένα audio formats χρησιμοποιούν τα δικά τους tag containers:
 - Vorbis Comment
 - Vorbis, FLAC, Theora, Speex, Opus
 - APE tags:
 - Monkey's Audio αλλά και MP3

Είδη Μουσικών Μεταδεδομένων

Περιγραφικά Μεταδεδομένα (descriptive metadata)

- για αναζήτηση, για οργάνωση, για ταξινόμηση μουσικής σε δομημένες μουσικές βιβλιοθήκες, για εκπομπή σε ραδιοφωνικούς σταθμούς, καθώς και για αναγνώριση ενός μουσικού έργου

Μεταδεδομένα ιδιοκτησίας (ownership metadata)

- εισήχθησαν για λόγους αναγνώρισης πνευματικών δικαιωμάτων αλλά και για λόγους προώθησης και αναγνωρισιμότητας των συντελεστών που συμμετέχουν σε μια μουσική παραγωγή (π.χ. ISRC αντίστοιχο του ISBN)

Μεταδεδομένα σύστασης (recommendation metadata)

- Περιγράφουν πώς ένα αρχείο ήχου «ακούγεται», π.χ. Genre, mood
- Συνοδεύουν τα αρχεία ήχου, αλλά δεν προσαρτώνται σε αυτά
- Αξιοποιούνται από αλγόριθμους σύστασης σε υπηρεσίες streaming (π.χ. Youtube, Spotify)
- Συνήθως είναι ιδιόκτητα και μπορεί να διαφέρουν από πάροχο σε πάροχο



Επεξεργασία μεταδεδομένων - MusicBrainz

MusicBrainz Picard

Add Folder Add Files Cluster Lookup Scan Lookup In Browser Save Info Remove Lookup CD Submit AcoustiDs Album [dropdown] [input] [search]

Title	Track No.	Length	Artist	Title
Regiment	3	4:11	Eno, Brian & Byrne, David	
Help Me Somebody	4	4:17	Eno, Brian & Byrne, David	
The Jezebel Spirit	5	4:57	Eno, Brian & Byrne, David	
Very, Very Hungry	6	3:21	Eno, Brian & Byrne, David	
Moonlight in Glory	7	4:30	Eno, Brian & Byrne, David	
The Carrier	8	4:20	Eno, Brian & Byrne, David	
A Secret Life	9	2:32	Eno, Brian & Byrne, David	

Tag	Original Value	New Value
Title	Very, Very Hungry	Very, Very Hungry
Artist	Eno, Brian & Byrne, David	Eno, Brian & Byrne, David
Album	My Life in the Bush of Ghosts	My Life in the Bush of Ghosts
Track Number	6	6
Length	3:21	3:21
encoder	Max 0.71	Max 0.71
encoding	FLAC settings: exhaustiveModelSearch:0 midSideStereo:1 I...	FLAC settings: exhaustiveModelSearch:0 midSideSt...
Genre	Ambient	Ambient
ISRC	USWB10102669	USWB10102669
mcn	0075597989427	0075597989427
Total Tracks	18	18

Show more details

0:47 [progress bar] 3:21 [volume icon] 1.0 x

Play Pause

06 Very, Very Hungry.flac

50%

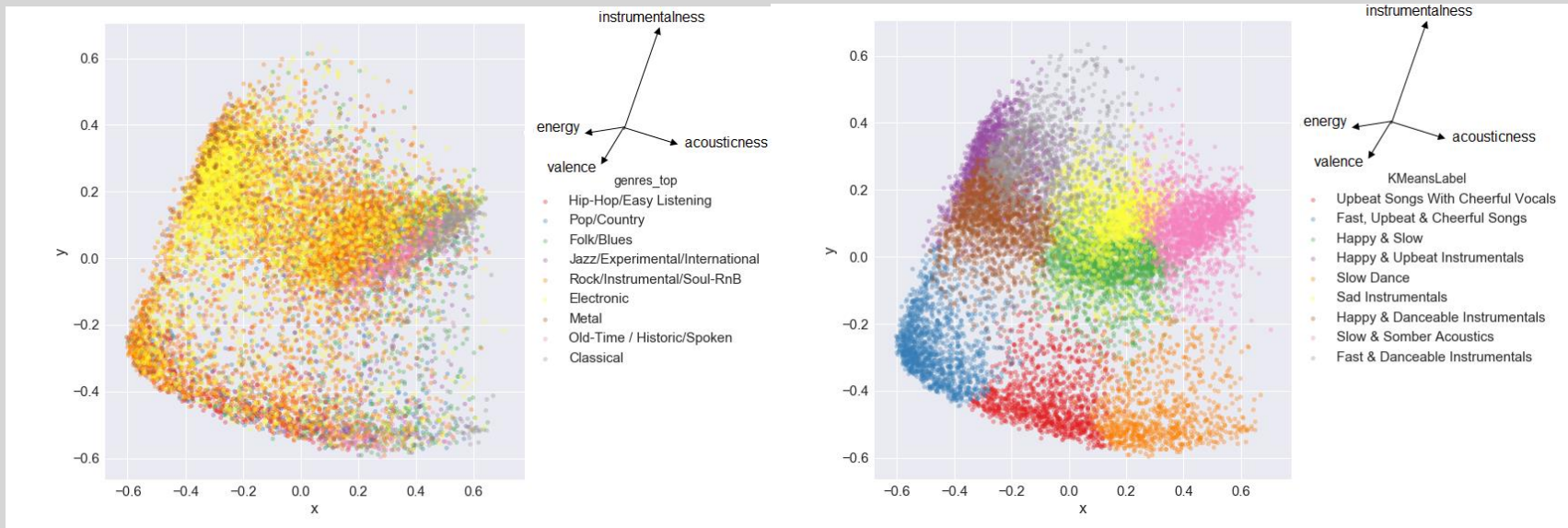
18 [album icon] 0 [CD icon] 0 [album icon] 0 [download icon] Listening on port 8000

Recommendation Metadata

- ❖ Περιγράφουν το πώς «ακούγεται» στους χρήστες ένα μουσικό έργο (π.χ. μουσικό είδος, ύφος, συναίσθημα κκ.)
- ❖ Λαμβάνονται ως εξής:
 1. Προστίθενται από χρήστες/ ακροατές ενός μουσικού έργου
 - User Tags, π.χ. #funky, #sad
 2. Παράγονται αυτόματα από αλγόριθμους συσταδοποίησης (clustering) και ταξινόμησης (classification), που εφαρμόζονται:
 - Είτε στα ίδια τα ηχητικά αρχεία,
 - π.χ. δημιουργία ομάδων ανά μουσικό είδος/ύφος
 - Είτε στα προφίλ χρηστών
 - π.χ. χρήστες παρόμοιας ηλικίας και τρόπου διαμονής τείνουν να προτιμούν μουσική συγκεκριμένου ύφους

Τι είναι η συσταδοποίηση - Clustering

❖ Ο αυτόματος διαχωρισμός οντοτήτων (π.χ. μουσικών αρχείων) σε ομάδες, με βάση τις ομοιότητές τους ως προς κάποιο/α χαρακτηριστικό/α

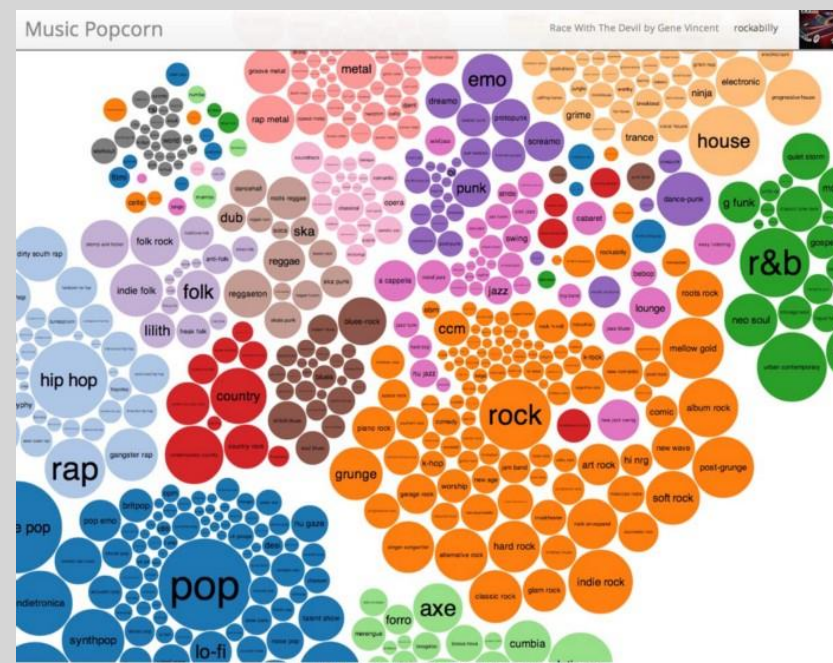


❖ <https://medium.com/latinxinai/discovering-descriptive-music-genres-using-k-means-clustering-d19bdea5e443>

Τι είναι η ταξινόμηση - Classification

❖ Η εύρεση της τάξης (class) στην οποία ανήκει μια οντότητα (π.χ. μουσικό αρχείο) από ένα πεπερασμένο πλήθος τάξεων:

➤ Π.χ. Genre classification



Classification vs. Clustering

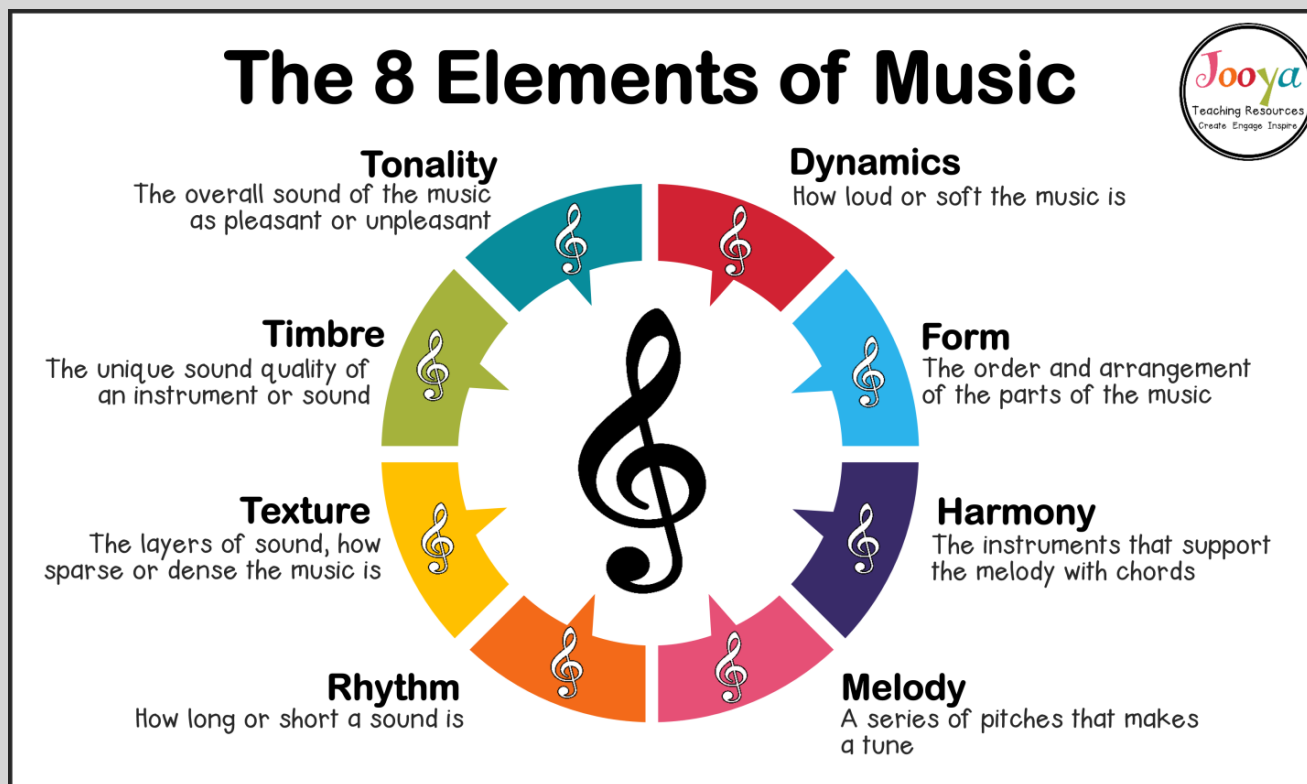
- ❖ Και τα δύο αφορούν ομάδες
- ❖ Στην επιστήμη δεδομένων και τη μηχανική μάθηση
 - Το classification επιτυγχάνεται με μεθόδους
 - επιβλεπόμενης μάθησης (supervised learning)
 - Πεπρασμένος αριθμός τάξεων, π.χ. 20 διαφορετικά genres
 - Χρειάζεται να διατίθενται επισημειωμένα δεδομένα, δηλ. ποιά οντότητα ανήκει σε ποιά κλάση
 - Το clustering γίνεται με μεθόδους **μη επιβλεπόμενης μάθησης (unsupervised learning)**
 - Βάσει χαρακτηριστικών ομοιότητας



Περιγραφή Μουσικού Περιεχομένου

Music Content

- Music has many different fundamentals or elements. Depending on the definition of "element" being used, these can include **pitch, beat or pulse, tempo, rhythm, melody, harmony, texture, style, allocation of voices, timbre or color, dynamics, expression, articulation, form, and structure.**



Παράδειγμα Περιγραφής Μουσικού Περιεχομένου

- ❖ Πρόκειται για πληροφορία η οποία η οποία μεταβάλλεται κατά τη διάρκεια ενός μουσικού έργου, όπως οι πληροφορίες που αναπαρίστανται σε μια παρτιτούρα:
 - Πότε αναπαράγεται η κάθε νότα/συγχορδία
 - Με τι τονικό ύψος, ένταση, χροιά (μουσικό όργανο)
 - Ποιες χρονικές στιγμές αντιστοιχούν σε χρονικά beats και ποιο είναι το tempo του κομματιού?
 - Ποιος είναι ο ρυθμός του κομματιού?

Σχετιζόμενες Τεχνολογίες

Πρότυπα μορφοποίησης (formats)

- File formats τα οποία υποστηρίζουν την αποθήκευση πληροφοριών περιγραφής μουσικού περιεχομένου

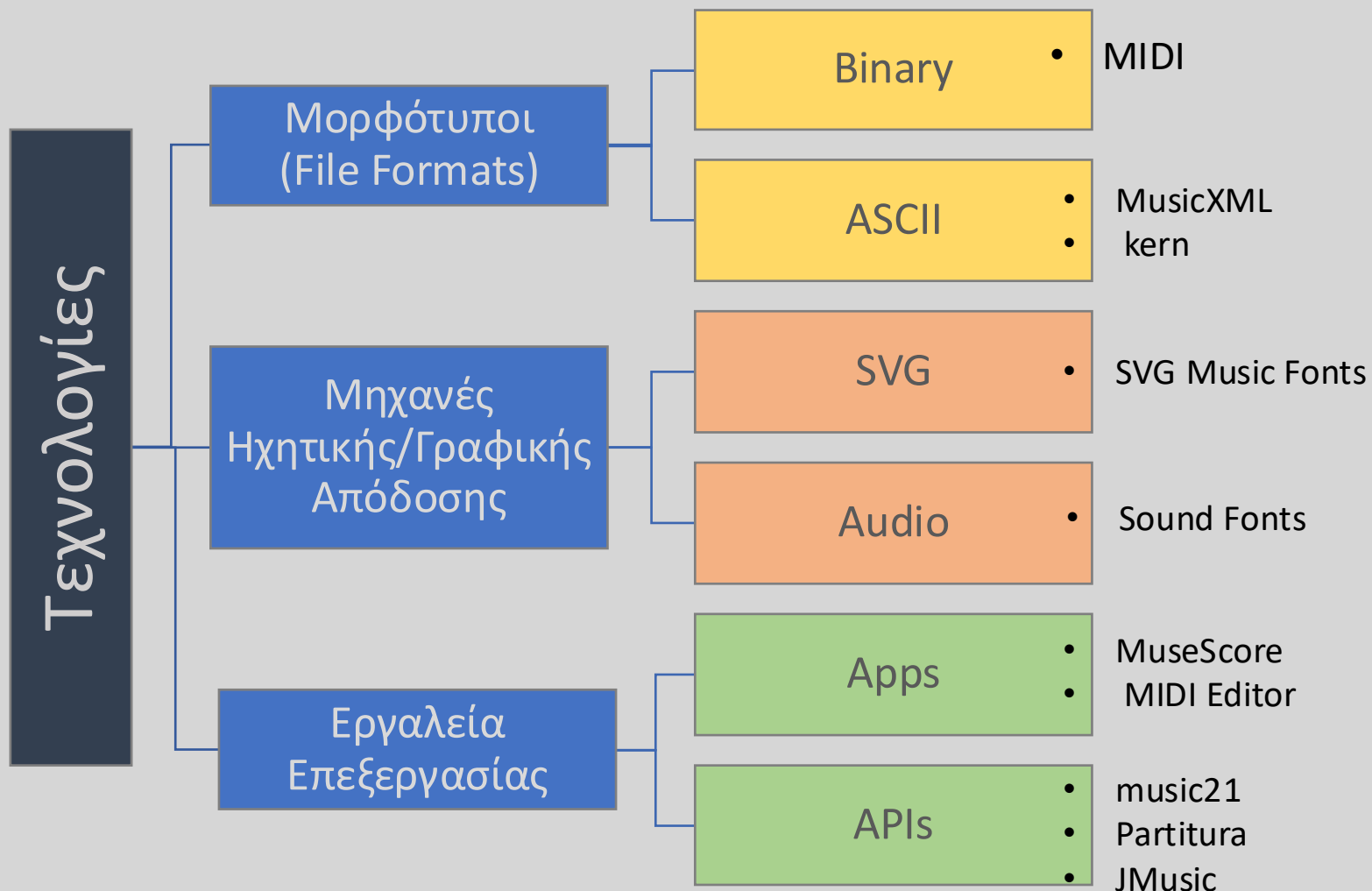
Εργαλεία Γραφικής/Ηχητικής Απόδοσης (rendering engines)

- Εφαρμογές που μπορούν να μετατρέψουν τα παραπάνω αρχεία είτε σε γραφική/σημειογραφική μορφή (παρτιτούρα) είτε σε ήχο

Εργαλεία ανάλυσης και επεξεργασίας (tools)

- Εργαλεία που συνθέτουν, αναλύουν, επεξεργάζονται τα αρχεία περιγραφής μουσικού περιεχομένου

Επισκόπηση σχετιζόμενων τεχνολογιών



Ερωτήματα

- ❖ Πώς η περιγραφή μουσικού περιεχομένου μετασχηματίζεται σε ήχο?
- ❖ Πως ο ήχος μετασχηματίζεται σε περιγραφή?
- ❖ Είναι όλες οι περιγραφές μουσικού περιεχομένου ισοδύναμες?