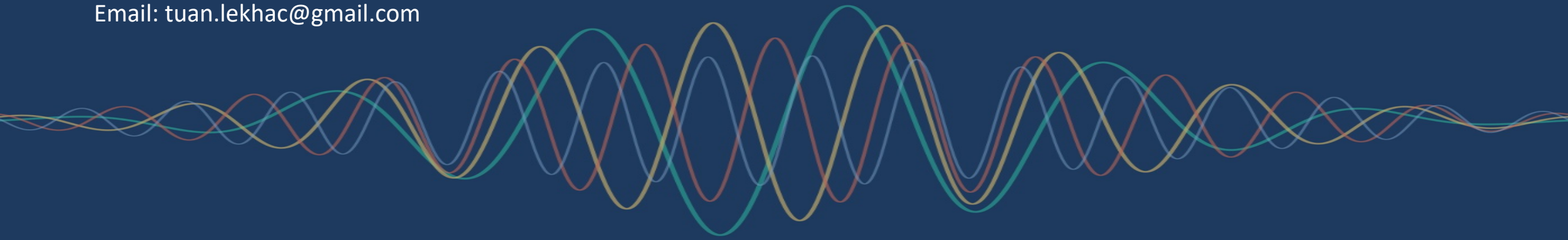
Khoa Công nghệ thông tin – Bộ môn Kỹ thuật máy tính • Ngành Công nghệ Đa phương tiện • Mã HP 531833

LẬP TRÌNH ÂM THANH

Buổi 1 – Tổng quan

Giảng viên: Lê Khắc Tuấn
Phone: 0962.11.05.1982
Email: tuan.lekhac@gmail.com



Ba câu hỏi lớn của buổi hôm nay

1

Môn học này là gì?

Lập trình + Âm thanh = ?
Kỹ thuật hay nghệ thuật?

2

Học để làm gì?

Ứng dụng thật, dự án thật,
nghề nghiệp thật

3

Học như thế nào?

Trong thời đại AI viết code
chỉ trong vài giây

Tiến trình buổi học (150 phút)

0	Khởi động: “Tắt tiếng thế giới”	10’
1	Cuộc sống gắn liền hình ảnh và âm thanh	15’
2	Giải mã tên môn học	20’
3	**Từ lớp học đến dự án thật**	30’
	Giải lao	10’
4	**Code xưa – code nay – vibe coding**	25’
5	Bản đồ môn học	15’
6	Học như thế nào?	10’
7	Demo: nghe giống – nhìn khác	5’
8	Tổng kết và nhiệm vụ về nhà	10’

Khởi động: “Tắt tiếng thế giới”

PHẦN 0

1

Xem phim KHÔNG có tiếng

Chuyện gì đang xảy ra?
Không khí thế nào?

2

Nghe tiếng KHÔNG có hình

Nhắm mắt lại.
Đoán lại chuyện gì đang xảy ra.

3

Xem lại đầy đủ

So sánh: lần nào em đoán
đúng hơn? Lần nào khó chịu hơn?

[Chèn đoạn phim 30–45 giây có tiếng mưa, xe cộ, nhạc nền, lời thoại]



Mất hình hay mất tiếng – em thấy khó chịu hơn?

Vì sao? Chỉ nghe tiếng, em đoán được những gì?

“Khán giả tha thứ cho hình xấu, nhưng không tha thứ cho tiếng xấu.”

Âm thanh mang thông tin, cảm xúc và nhịp – một nửa trải nghiệm đa phương tiện

PHẦN 1 • BRAINSTORM



Từ sáng đến giờ, em đã nghe bao nhiêu âm thanh do MÁY tạo ra hoặc xử lý?

2 phút – mỗi người tự liệt kê, sau đó cả lớp gom nhóm

Bốn nhóm âm thanh trong một ngày

PHẦN 1



Câu hỏi: trong số này, âm thanh nào KHÔNG đi qua một đoạn code nào?

Âm thanh nào KHÔNG đi qua code?

PHẦN 1

Không qua code

Tiếng chim hót, tiếng người nói chuyện trực tiếp, tiếng mưa ngoài cửa sổ…

…nhưng ngay khi được ghi âm, gọi điện, phát qua loa Bluetooth, chúng đã đi qua code: nén, giải nén, lọc ồn, khử vọng.

Code vô hình: khử tiếng vọng

Khi gọi video, vì sao em không nghe thấy giọng của chính mình vọng lại từ loa của người kia?

→ Thuật toán khử tiếng vọng (echo cancellation) chạy liên tục trên điện thoại của cả hai người.

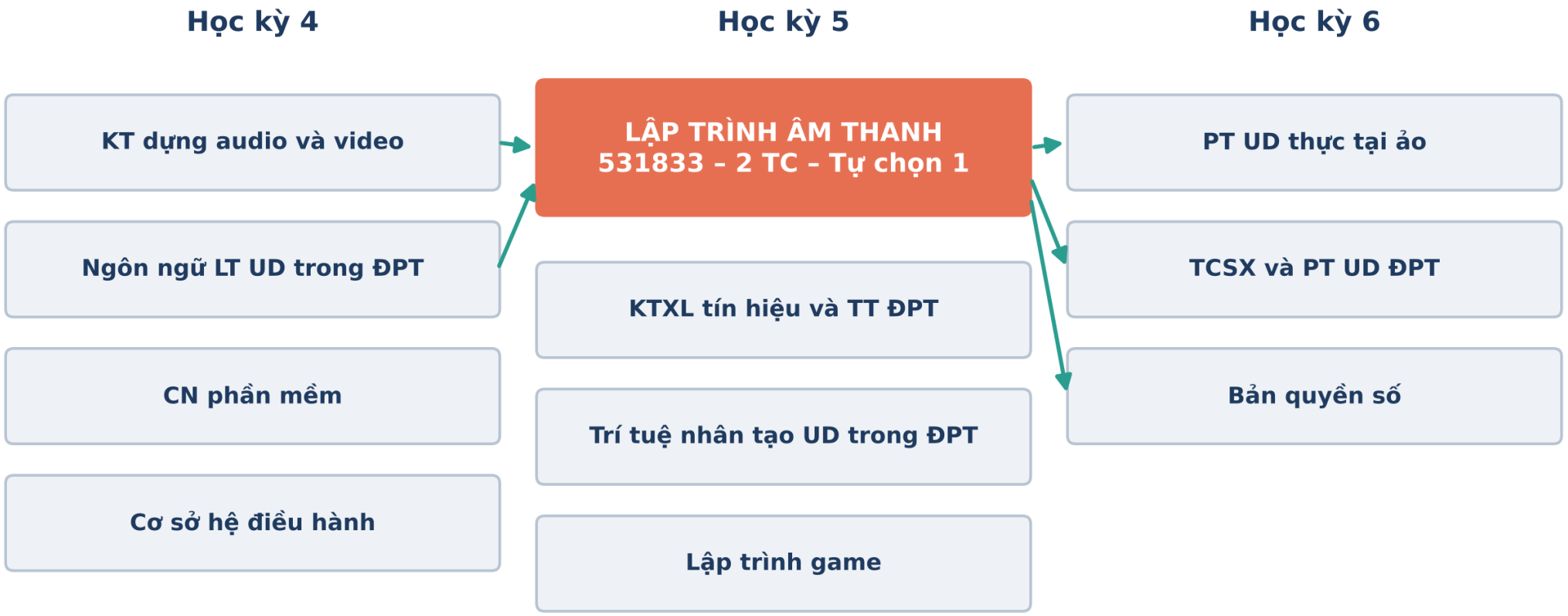


Chỉ dựa vào tên gọi, môn này học gì?

“Lập trình” + “Âm thanh” – kỹ thuật hay nghệ thuật?

Vị trí môn học trong chương trình đào tạo

PHẦN 2

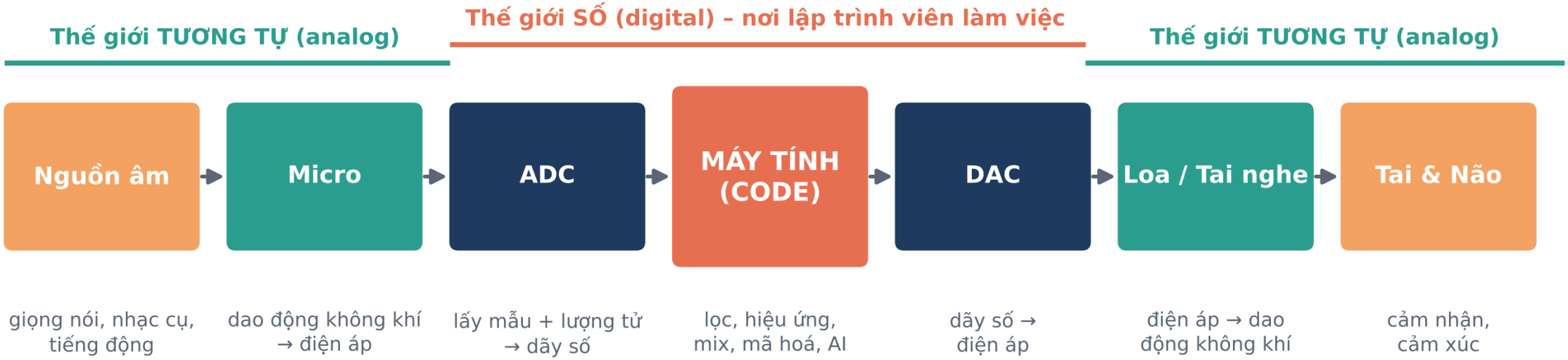


Môn học là điểm hội tụ: dựng audio (HK4) + lập trình + xử lý tín hiệu + AI → ứng dụng thực tại ảo, sản phẩm ĐPT, bản quyền số (HK6)

Nguồn: Kế hoạch giảng dạy CTĐT ngành CNTT/chuyên ngành CN Đa phương tiện (Khóa 69), ĐH Xây dựng Hà Nội

Chuỗi tín hiệu: lập trình viên đứng ở đâu?

PHẦN 2



Mọi thứ trong môn học nằm ở khối màu cam. Nhưng để code tốt, phải hiểu cả các khối hai bên: micro thu được gì, loa phát được gì, tai người cảm nhận thế nào.

Ba tầng kiến thức

PHẦN 2



Lập trình nối ba tầng

- Bộ môn KTMT mạnh ở tầng 1: phần cứng, mạng
- Ngành ĐPT hướng tới tầng 3: sản phẩm, nội dung
- Môn học này dạy tầng 2 và dùng code để nối cả ba

Môn học đứng ở đâu?

PHẦN 2

Kỹ sư DSP thuần

Toán sâu, thuật toán xử lý tín hiệu, ít làm sản phẩm

LẬP TRÌNH ÂM THANH

Hiểu nguyên lý + tự viết được công cụ cho sản phẩm đa phương tiện

Người bấm phần mềm

Dùng thành thạo công cụ có sẵn, khó đi xa hơn công cụ cho phép

Cầu nối giữa kỹ thuật và sản xuất nội dung



Học xong môn này, em có thể làm được gì?

Có kiếm tiền được từ nó không?

Case study 1: Số hóa di tích – web 360° có âm thanh

PHẦN 3



Người xem tham quan di tích qua trình duyệt, xoay góc nhìn 360°. Âm thanh giúp họ cảm thấy “đang đứng ở đó”.

1

Âm thanh nền

Gió, chim, lá cây, suối – phát lặp liên tục

2

Nguồn âm có vị trí

Chuông chùa, tiếng mõ – gắn với một điểm trong không gian

3

Thuyết minh

Giọng thật hoặc TTS, nhiều ngôn ngữ

4

Chuyển cảnh

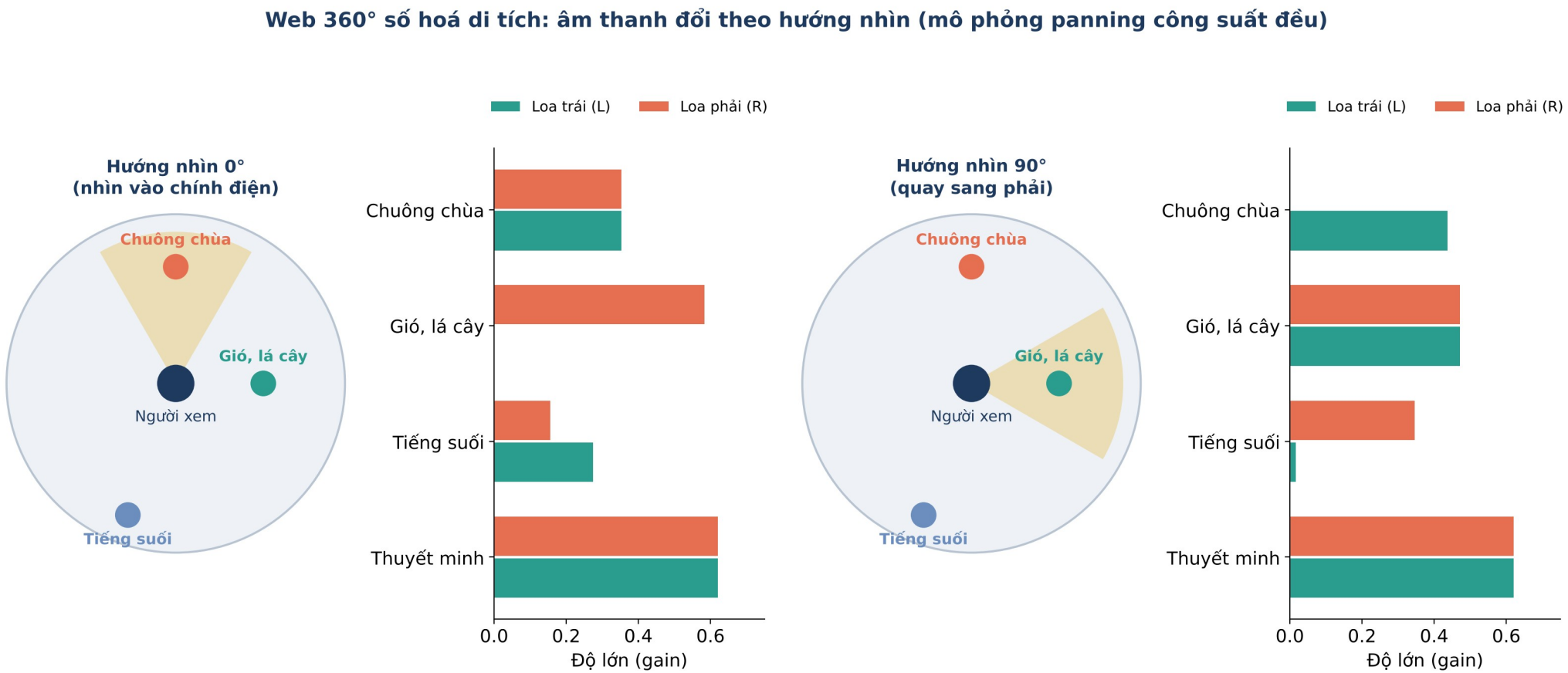
Crossfade: âm cũ nhỏ dần, âm mới to dần



DEMO: tắt âm thanh và xoay → bật âm thanh và xoay lại → chuyển điểm tham quan

Âm thanh không gian: đổi theo hướng nhìn

PHẦN 3



Nhìn thẳng vào chính điện: chuông ở giữa, hai loa bằng nhau. Quay sang phải 90°: chuông nằm bên trái nên loa trái to hơn hẳn. Giọng thuyết minh “gắn vào đầu” nên không đổi.

Kiến thức của môn đứng sau tour 360°

PHẦN 3

Ghi âm hiện trường, lọc ồn	Chương 4
Chuẩn hóa độ lớn giữa các điểm tham quan	Chương 4
Trộn nhiều lớp, crossfade khi chuyển cảnh	Chương 5
Web Audio API: PannerNode, GainNode, AudioListener	Chương 6
Nén file cho web để tải nhanh	Chương 2



Quay mặt từ cổng tam quan vào chính điện, tiếng chuông chùa nên thay đổi thế nào?

Làm sao để code "biết" em đang nhìn hướng nào?

Case study 2: Loa phường IP tích hợp TTS

PHẦN 3



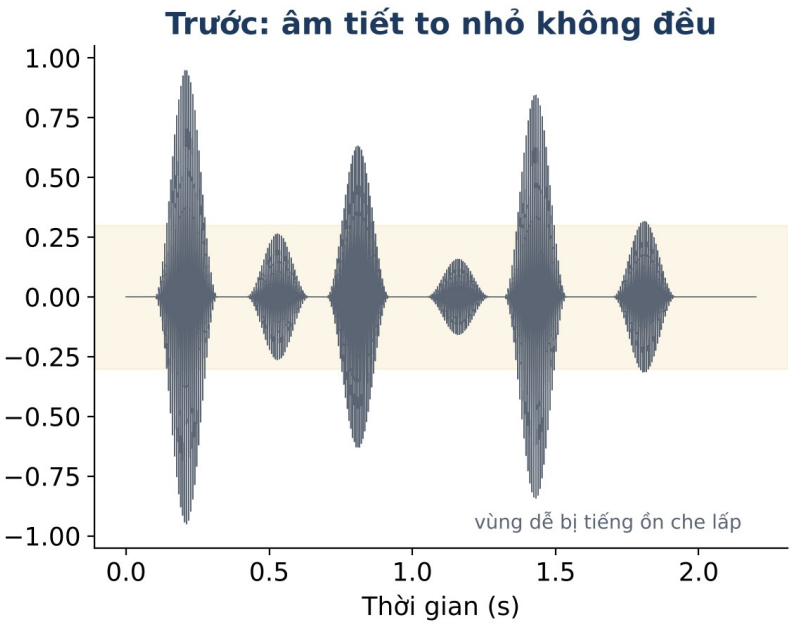
Bài toán thật cần giải:

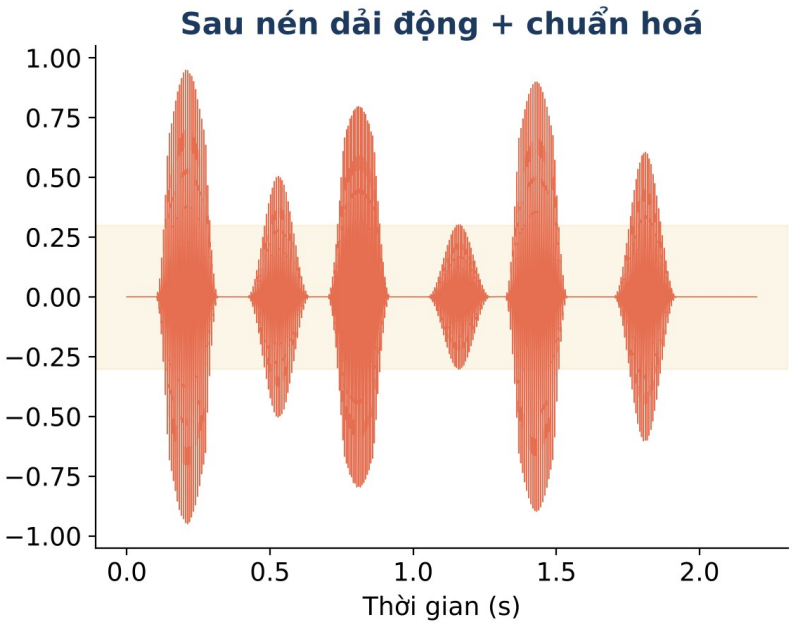
Nghe rõ lời ngoài trời, giữa tiếng xe cộ	Đọc đúng tên riêng, địa danh, số, ngày tháng	Giọng tự nhiên, hợp vùng miền	Phát đồng bộ nhiều cụm, độ trễ mạng

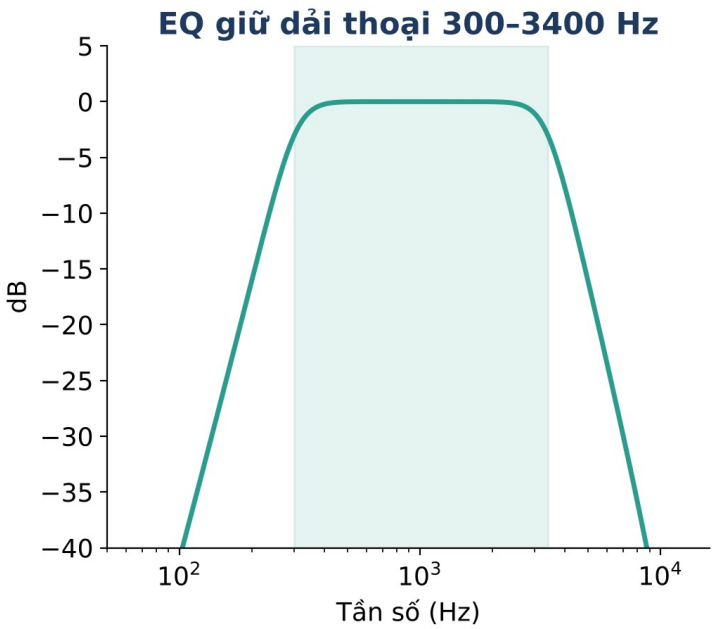
DEMO: gõ một thông báo → hệ thống đọc bằng TTS → phát tới cụm loa

Vì sao loa phường khó nghe rõ lời?

PHẦN 3







Loa nén chỉ phát tốt dải 300–3400 Hz
→ **EQ giữ dải thoại**

Tiếng xe cộ che lấp âm tiết nhỏ
→ **nén dải động**

Các bản tin to nhỏ không đều
→ **chuẩn hóa độ lớn**



Em sẽ xử lý tín hiệu thế nào để loa phường nghe rõ lời?

Và: TTS đọc “Thứ 2, ngày 10/11” thì có thể sai thế nào? Sửa ở khâu nào?

Bản đồ nghề nghiệp: đứng ở vùng giữa

PHẦN 3

Kỹ thuật sâu

Số lượng ít, chủ yếu ở các công ty lớn: codec, DSP, AI giọng nói

Vùng giữa

Số hóa di tích, loa IP, app âm thanh, game audio – cần người hiểu cả kỹ thuật và sản phẩm

Sản xuất nội dung

Số lượng nhiều: podcast, video, MV, quảng cáo, mạng xã hội

Ứng dụng khác: chuẩn hóa loudness cho YouTube/Spotify • app thư giãn, hỗ trợ giấc ngủ • nhận dạng giọng nói • sinh nhạc bằng AI • kiểm định chất lượng, bản quyền âm thanh

“Những gì các em học trong 15 buổi tới, thầy đang dùng để làm sản phẩm thật, cho khách hàng thật.”



GIẢI LAO 10 PHÚT

Mở tour 360° trên điện thoại và tự trải nghiệm



Ai đã từng dùng AI để viết code?

Ai đã từng nộp bài có code do AI viết mà mình không hiểu hết?

Lập trình ngày xưa và ngày nay

PHẦN 4

NGÀY XƯA (trước 2022)		NGÀY NAY (vibe coding với AI)
Đọc sách, tài liệu API, tra Stack Overflow	→	Mô tả bằng lời → AI sinh code trong vài giây
Gõ từng dòng, nhớ cú pháp	→	Không cần nhớ cú pháp
Debug hàng giờ, đọc từng thông báo lỗi	→	AI gợi ý sửa lỗi, giải thích code
Rào cản vào nghề CAO	→	**Rào cản vào nghề THẤP**
Ai vượt qua được thì HIỂU SÂU	→	**Dễ có code chạy mà KHÔNG HIỂU**

Kỹ năng quan trọng đã đổi: từ “viết được code” sang “hiểu, kiểm chứng và sửa được code”

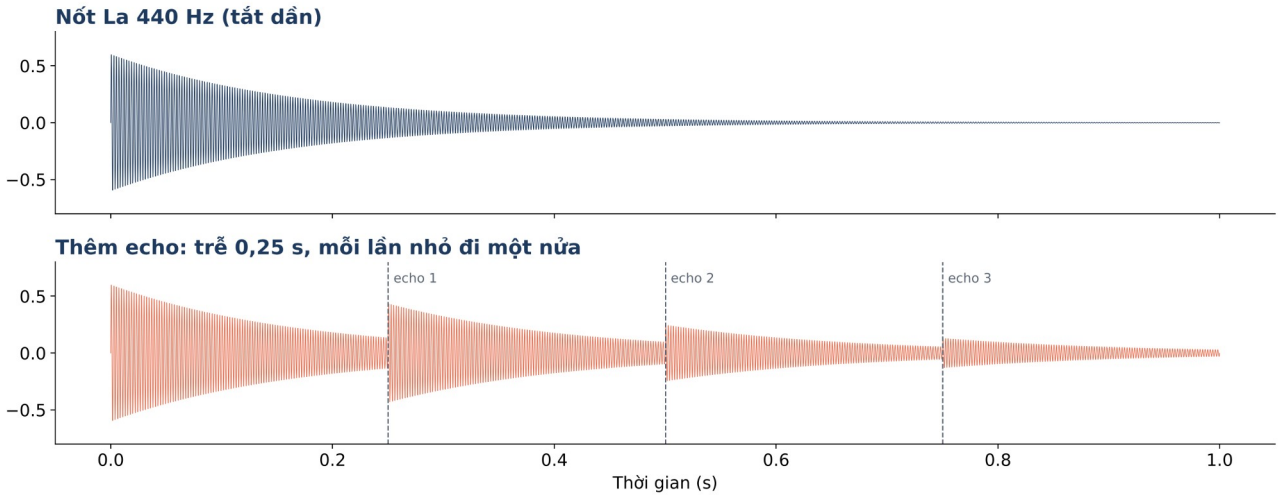
Demo trực tiếp: nhờ AI viết code tạo echo

PHẦN 4

Prompt

"Viết chương trình Python tạo nốt La 440 Hz dài 1 giây, thêm hiệu ứng echo trễ 0,25 giây, lưu ra file WAV."

```
tre = int(0.25 * fs)
for k in range(1, 4):
    co_echo[k*tre:] += (0.5**k) * \
        not_la[:len(not_la) - k*tre]
```



Code chạy được ngay. Nhưng…

Code chạy rồi. Còn em thì sao?

PHẦN 4



Giải thích được từng dòng không?

Vì sao nhân 2π? Vì sao 0,25 giây lại thành 11 025?



Sửa được khi lỗi không?

File kêu rè, vỡ tiếng – em hỏi lại AI hay tự tìm nguyên nhân?



Phát hiện được AI sai không?

Code vẫn chạy nhưng âm thanh sai nhẹ – em có nghe ra không?

Ba rủi ro của vibe coding

PHẦN 4

1

Ảo tưởng năng lực

Tưởng mình biết vì code chạy được

2

Nợ kỹ thuật

Code chắp vá, không ai hiểu để sửa

3

AI sai mà không biết

Lỗi âm thanh nhẹ rất khó phát hiện nếu thiếu kiến thức

Mặt mạnh: nhanh, ai cũng tiếp cận được. Môn học không cấm AI – môn học đặt ra luật chơi.

Luật chơi của môn học với AI

PHẦN 4

VÙNG PHẢI TỰ CODE

- ✓ Thuật toán DSP cốt lõi: lọc, FFT, nén
- ✓ Hiệu ứng: echo, reverb, EQ
- ✓ Phân tích & đo đạc tín hiệu
- ✓ Code dùng để bảo vệ bài tập lớn

VÙNG ĐƯỢC DÙNG AI

- ✓ Giải thích code, giải thích lỗi
- ✓ Review và gợi ý cải tiến
- ✓ Giao diện (HTML/CSS), tài liệu
- ✓ Tìm ý tưởng, lên kế hoạch

1 KHAI BÁO
đã dùng AI ở đâu, prompt gì

2 GIẢI THÍCH
được từng dòng code đã nộp

3 BẢO VỆ
trả lời câu hỏi, sửa code tại chỗ

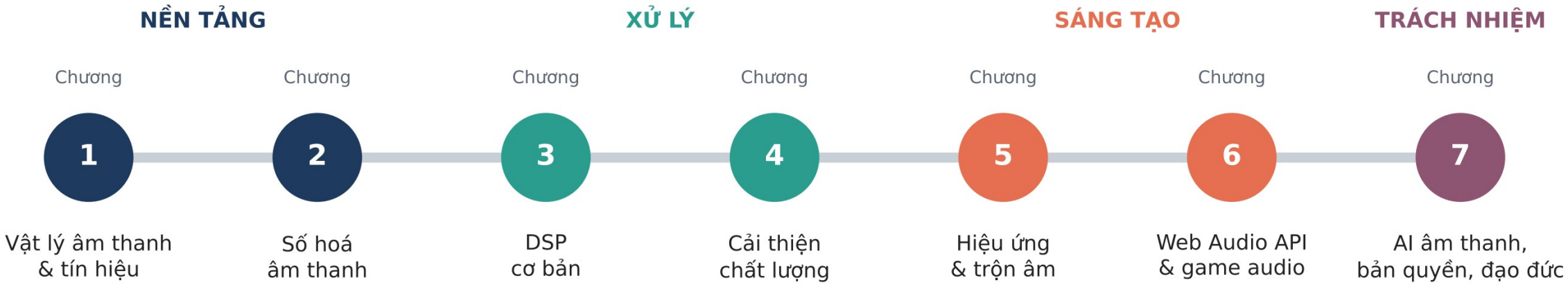


“Hiểu code” quan trọng hơn “có code”

AI là trợ lý, không phải người làm thay

Lộ trình: 7 chương, 4 giai đoạn bài tập lớn

PHẦN 5



Mỗi chương gắn với dự án thật

PHẦN 5

Chương	Nội dung	Dùng ở đâu trong dự án thật
1	Vật lý âm thanh, tín hiệu tương tự	Micro, loa, amply của hệ thống loa IP
2	Số hóa: lấy mẫu, lượng tử, định dạng	Định dạng nén cho web 360° và luồng loa IP
3	DSP cơ bản: phổ, FFT, bộ lọc	Phân tích tiếng ồn hiện trường
4	Lọc ồn, chuẩn hóa, nén dải động	Độ rõ lời loa IP; âm lượng đều giữa các điểm
5	Hiệu ứng, trộn âm, crossfade	Trộn lớp âm thanh, chuyển cảnh tour 360°
6	Web Audio API, game audio	Âm thanh không gian theo hướng nhìn
7	AI âm thanh, bản quyền, đạo đức	TTS cho loa IP; bản quyền nhạc nền, giọng đọc

Công cụ, bài tập lớn và đánh giá

PHẦN 5

Công cụ

- Python: numpy, scipy, soundfile, librosa, pyloudnorm
- JavaScript: Web Audio API
- Audacity: nghe, chỉnh sửa
- Spek / script: xem phổ
- ffmpeg: chuyển định dạng

Bài tập lớn

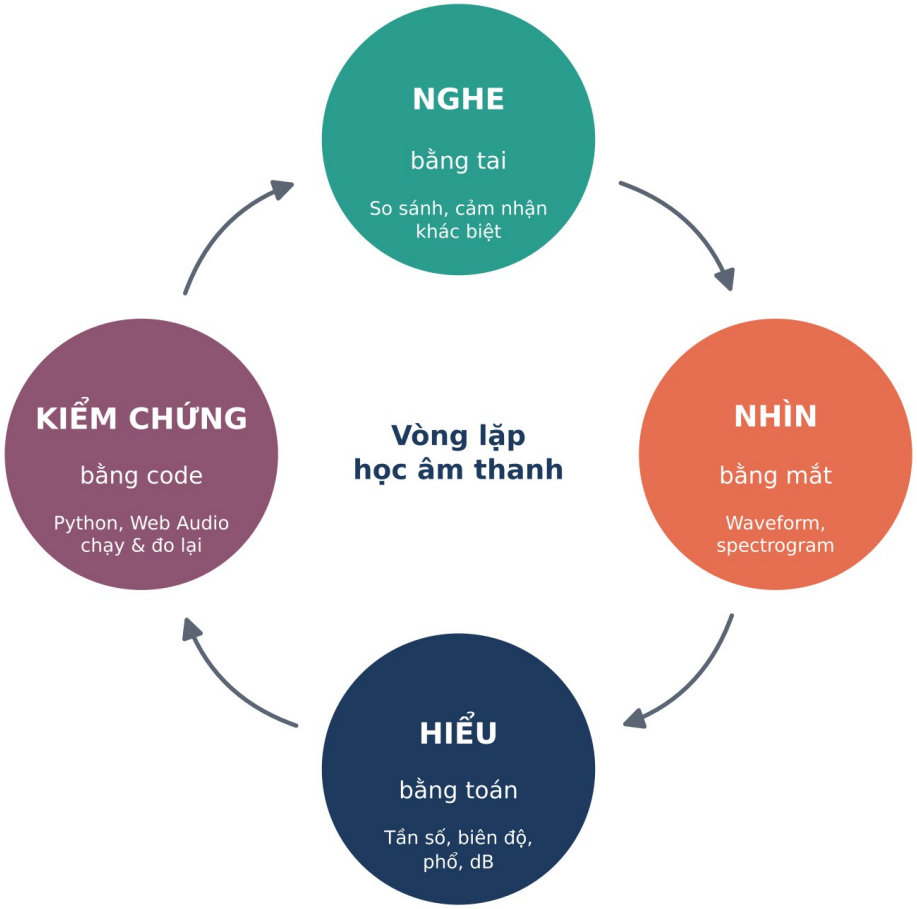
- Làm theo nhóm, 4 giai đoạn GĐ1–GĐ4
- Sản phẩm cuối: một ứng dụng web âm thanh hoàn chỉnh
- Tham chiếu: app “Dịu” – thở 4-7-8 với nhiều lớp âm thanh
- Được đề xuất đề tài gắn với dự án thật

Đánh giá

- Chuyên cần: …%
- Giữa kỳ: …%
- Bài tập lớn: …%
- Cuối kỳ: …%
- Bảo vệ code là bắt buộc

Học âm thanh như thế nào?

PHẦN 6

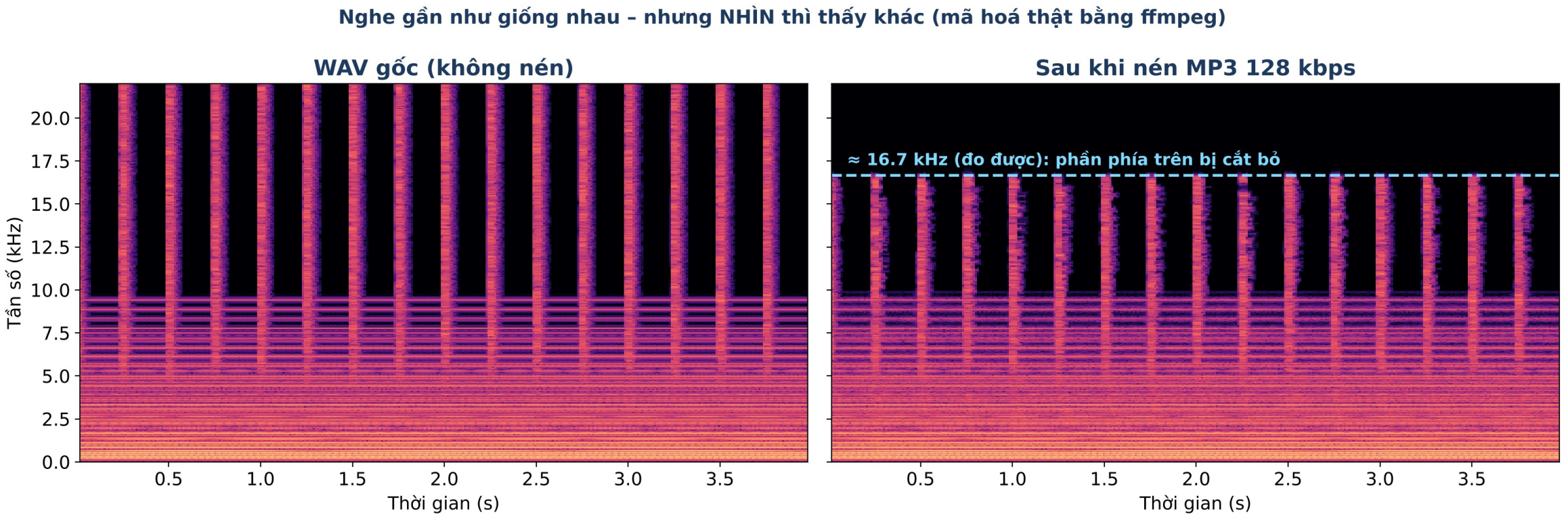


Âm thanh vô hình – không thể chỉ học bằng mắt đọc sách.

- Chuẩn bị: tai nghe CÓ DÂY (Bluetooth có độ trễ và tự nén), laptop đã cài Python
- Làm việc nhóm: chia nhóm từ buổi 2, người mạnh code + người mạnh nội dung
- Thái độ: thử nghiệm, không sợ sai – lỗi nghe được là bài học đáng nhớ

Nghe giống nhau – nhìn thì khác

PHẦN 7



MP3 128 kbps cắt bỏ phần phía trên khoảng 16,7 kHz. Cách này còn dùng để phát hiện file “lossless giả”. Tai có thể bị đánh lừa – phổ thì không.

Trả lời ba câu hỏi lớn

PHẦN 8

Là gì?

Dùng lập trình can thiệp vào khối giữa của chuỗi tín hiệu âm thanh – cầu nối kỹ thuật và nội dung

Để làm gì?

Làm sản phẩm thật: tour 360° có âm thanh không gian, loa phường IP có TTS, app, game, podcast…

Học thế nào?

Nghe – nhìn – hiểu – kiểm chứng; dùng AI có kiểm soát, khai báo và bảo vệ được code

Nhiệm vụ về nhà

PHẦN 8

1 Cài môi trường

Python 3, numpy, scipy, soundfile, matplotlib, Audacity. Chạy demo1_not_la_echo.py, chụp màn hình.

2 Ghi âm 10 giây

Âm thanh quanh nơi em ở. Mở bằng Audacity, quan sát dạng sóng.

3 Viết 3 câu trả lời

“Em học môn này để làm gì?”

4 Chọn dự án

Nếu được tham gia số hóa di tích hoặc loa phường IP, em muốn làm phần nào? Vì sao?

Phiếu khảo sát đầu vào

PHẦN 8

[Chèn mã QR
Google Form]

- Em đã biết Python ở mức nào?
- Em có chơi nhạc cụ, hát, làm nhạc, làm video không?
- Em đã dùng AI để viết code chưa? Công cụ nào?
- Em có laptop và tai nghe có dây không?
- Em mong đợi điều gì nhất ở môn học này?

Hẹn gặp lại ở Buổi 2

Mang theo laptop và tai nghe có dây!

