MỤC LỤC

Chương 12. Trí tuệ nhân tạo và công nghệ mới 8
12.1 Vị trí của chương trong cuốn sách 9
12.2 Ba loại tác vụ mà máy làm được 12
12.3 Ba mức can thiệp vào một khâu công việc 14
12.4 AI xen vào đâu trong dây chuyền 15
12.5 Hai số đo của một công cụ tạo sinh 17
12.6 Sáu câu hỏi để đánh giá một công cụ 18
12.7 Bài thực hành số 1 — đo tỉ lệ dùng được 20
12.8 Tiền kỳ 21
12.9 Trường quay 23
12.10 Hậu kỳ hình — nhận dạng và biến đổi 24
12.11 Hậu kỳ hình — tạo sinh 26
12.12 Hậu kỳ tiếng — dọn tiếng và tách nguồn 29
12.13 Hậu kỳ tiếng — giọng nhân bản và lồng tiếng 32
12.14 Quyết định dùng hay không — và bài thực hành số 2 33
12.15 Sản xuất ảo — phần dựng cảnh thời gian thực 36
12.16 ICVFX — quy trình và ê-kíp mới 38
12.17 XR và các phương tiện mới — đọc đúng hiện trạng 41
12.18 Hạ tầng — sản xuất trên nền IP và trên đám mây 43
12.19 Hướng nén mới dùng mạng thần kinh 45
12.20 Bài thực hành số 3 — kịch bản kỹ thuật cho một cảnh sản xuất ảo 50
12.21 Luật Trí tuệ nhân tạo của Việt Nam 51
12.22 Điều 11 và nghĩa vụ gắn nhãn 53
12.23 Nghĩa vụ tương ứng ở thị trường ngoài nước 55
12.24 Xuất xứ nội dung: công cụ và giới hạn thật 56
12.25 Quyền hình ảnh và giọng nói — bản sao số của người biểu diễn 58
12.26 Quyền tác giả đối với nội dung có AI tham gia 60
12.27 Đạo đức nghề và cách đọc một dự báo công nghệ 61
12.28 Bài thực hành số 4, bảng tra nhanh và tổng kết 63
12.29 Câu hỏi và bài tập ôn tập 66

DANH MỤC HÌNH VẼ

Hình 12.1 Bản đồ Chương 12 — bốn phần và câu hỏi trung tâm của mỗi phần.....10
Hình 12.2 Ba bậc tuổi thọ của vật liệu viết, và tỉ lệ mà chương này dành cho từng bậc 11
Hình 12.3 Ba loại tác vụ: nhận dạng, biến đổi và tạo sinh — phân theo quan hệ giữa đầu vào và đầu ra..... 12
Hình 12.4 Tính không lặp lại của tác vụ tạo sinh: cùng một mô tả, mỗi lần chạy ra một kết quả khác..... 14
Hình 12.5 Ba mức can thiệp: trợ giúp, thay một khâu, tạo sinh — và trách nhiệm tương ứng của con người ở mỗi mức..... 14
Hình 12.6 Chín trạm của dây chuyền và mức can thiệp của máy ở từng trạm.....16
Hình 12.7 Ba mốc bàn giao và lý do chúng không giao được cho máy.....17
Hình 12.8 Tỉ lệ kết quả dùng được thay đổi thế nào theo độ khó của việc.....17
Hình 12.9 Sáu câu hỏi để đánh giá một công cụ, và thứ tự hỏi.....19
Hình 12.10 Bài thực hành số 1 — cấu trúc bài và kết quả cần có.....20
Hình 12.11 Sáu khâu của tiền kỳ và mức can thiệp của máy ở từng khâu.....21
Hình 12.12 Bẫy trung bình: mô hình kéo mọi phương án về phía tâm của dữ liệu huấn luyện.....22
Hình 12.13 Bốn ứng dụng ở trường quay, tất cả đều ở mức trợ giúp.....23
Hình 12.14 Phiên âm tự động trong bảng Text của Adobe Premiere Pro, từ một đoạn thoại tiếng Anh do giọng tổng hợp đọc; có lỗi nhận dạng cần kiểm lại.....24
Hình 12.15 Sáu ứng dụng hậu kỳ hình thuộc nhóm nhận dạng và biến đổi.....25
Hình 12.16 Generative Fill trong Photoshop trước khi tạo sinh: vùng chọn và ô nhập mô tả (tuỳ chọn). Ảnh nền là ảnh cầu vồng trên núi từ Wikimedia Commons 26
Hình 12.17 Ảnh do mô hình Stable Diffusion 3.5 tạo ra từ câu lệnh văn bản (tác giả gọi là synthograph); không phải ảnh chụp.....27
Hình 12.18 Bốn ứng dụng tạo sinh ở hậu kỳ hình và giới hạn của từng ứng dụng.....27
Hình 12.19 Nâng độ phân giải: chi tiết ở bản nâng cấp là chi tiết được bịa ra, không phải chi tiết được tìm lại..... 29
Hình 12.20 Enhance Speech trong Essential Sound của Adobe Premiere Pro: dọn tiếng nói bằng một nút bấm (tệp thoại mẫu tổng hợp).....30
Hình 12.21 Hộp thoại Noise Reduction (giảm tạp âm theo dấu vân tay) của Adobe Audition, áp dụng cho tệp thoại có tạp âm mẫu tổng hợp.....30
Hình 12.22 Năm việc dọn tiếng và mức cải thiện so với công cụ truyền thống.....31
Hình 12.23 Phổ tần của tệp thoại có tạp âm hồng mẫu tổng hợp, trước (trái) và sau (phải) khi giảm tạp âm trong Adobe Audition.....32
Hình 12.24 Bốn năng lực của nhóm giọng nhân bản, xếp theo mức rủi ro pháp lý.....32
Hình 12.25 Hai đường chi phí và điểm hoà vốn, vẽ cho ba giá trị của tỉ lệ dùng được. 34
Hình 12.26 Bảng quyết định: bốn tổ hợp của tỉ lệ dùng được và số lần phải làm.....35
Hình 12.27 Toàn bộ Phần B đặt lên hai trục: ba loại tác vụ và ba mức can thiệp.....35
Hình 12.28 Bài thực hành số 2 — bốn bước và kết quả cần có.....36
Hình 12.29 Trường quay sản xuất ảo cỡ nhỏ: tường LED cong hiển thị cảnh nền dựng thời gian thực, giàn treo đèn phía trên (Đại học Portsmouth).....37
Hình 12.30 Một khung hình sản xuất ảo được dựng như thế nào — ba nguồn đầu vào và hai vùng trên tường..... 37
Hình 12.31 Bộ sản xuất ảo trưng bày tại triển lãm BIRTV 2025: cảnh nền trên tường LED, cánh tay robot mang máy quay và bàn giám sát.....38
Hình 12.32 Bốn cái được của ICVFX và cái giá ít người nói trước.....39

Hình 12.33 Khối lượng công việc dời từ hậu kỳ sang tiền kỳ.....39
Hình 12.34 Ba vai trò mà ê-kíp truyền thống không có.....40
Hình 12.35 Bàn điều khiển và vùng quay của một trường quay sản xuất ảo: người vận hành ngồi tại bàn, máy quay và tường LED ở phía sau.....41
Hình 12.36 Thị trường thiết bị đội đầu rẽ hướng: sản lượng một sản phẩm cao cấp và cơ cấu doanh thu năm 2025.....42
Hình 12.37 Sáu câu hỏi để đọc một phương tiện mới.....43
Hình 12.38 Hạ tầng lai ghép: phần đã chuyển sang IP, phần vẫn chạy SDI, và ba điều kiện để chuyển.....43
Hình 12.39 Ngân sách độ trễ của một buổi dựng từ xa, và ba nhóm công việc theo mức đưa được lên đám mây.....44
Hình 12.40 Dãy tủ máy chủ của một siêu máy tính (Pleiades) trong trung tâm tính toán 45
Hình 12.41 Adobe Media Encoder đặt AAC (bộ mã hoá cổ điển, thiết kế thủ công) ở 24 kbit/s, đơn kênh, cho một đoạn thoại mẫu tổng hợp; đây là mốc đối chiếu với nén bằng mạng thần kinh.....46
Hình 12.42 Hai đường mã hoá và mức nén đạt được ở tiếng nói.....47
Hình 12.43 MediaInfo: đoạn thoại mẫu tổng hợp ở dạng PCM 1 152 kbit/s (trên) và sau khi nén bằng bộ mã hoá cổ điển AAC LC ở 24 kbit/s (dưới).....48
Hình 12.44 Ba điều kiện để một bộ mã hoá mới thật sự được dùng.....49
Hình 12.45 Bài thực hành số 3 — đề bài, bốn phần việc và kết quả cần có.....50
Hình 12.46 Ba mức rủi ro theo Luật số 134/2025/QH15 và ô của nghề sản xuất nội dung.....53
Hình 12.47 Điều 11 — bốn nghĩa vụ minh bạch và điều khoản dành riêng cho tác phẩm điện ảnh, nghệ thuật.....53
Hình 12.48 Sơ đồ ba câu hỏi: có phải gắn nhãn không, và gắn ở mức nào.....55
Hình 12.49 Hai hệ thống pháp luật độc lập với yêu cầu gần như giống nhau.....55
Hình 12.50 Bản kê xuất xứ đi theo tệp qua bốn chặng, và ba giới hạn phải nói thẳng. .56
Hình 12.51 Chuỗi xuất xứ đứt ở đâu — hai chỗ đứt phổ biến nhất.....57
Hình 12.52 Sáu điểm bắt buộc của một văn bản đồng ý cho bản sao số.....58
Hình 12.53 Ba câu hỏi quyền tác giả mà nhiều người hay gộp làm một.....60
Hình 12.54 Bốn ranh giới nghề nghiệp không nên vượt, dù luật chưa cấm.....61
Hình 12.55 Đối chiếu bảng dự báo với thực tế và ba quy luật rút ra được.....62
Hình 12.56 Bài thực hành số 4 — đề bài, bốn phần việc và cách tự đánh giá.....63
Hình 12.57 Bảng tra nhanh — làm gì trong từng tình huống.....64

DANH MỤC BẢNG BIỂU

Bảng 12.1 Bốn phần của chương.....9
Bảng 12.2 Ba mức can thiệp và điều thay đổi theo từng mức.....15
Bảng 12.3 Sáu câu hỏi, điều cần tìm và dấu hiệu đáng lo.....19
Bảng 12.4 Bốn ứng dụng ở trường quay.....23
Bảng 12.5 Sáu ứng dụng và mức chín muồi tới thời điểm biên soạn.....25
Bảng 12.6 Năm việc dọn tiếng.....31
Bảng 12.7 Cùng một công cụ, ba tỉ lệ dùng được, ba kết luận khác nhau.....34
Bảng 12.8 Hai văn bản cần biết.....52
Bảng 12.9 Bốn nghĩa vụ: ai phải làm và làm gì.....54
Bảng 12.10 Sáu điểm bắt buộc và lỗi thường gặp.....59
Bảng 12.11 Đối chiếu một số dự báo tiêu biểu với thực tế.....62

DANH MỤC KÝ HIỆU VÀ CHỮ VIẾT TẮT

Ký hiệu viết tắt	Thuật ngữ tiếng Anh đầy đủ	Ý nghĩa chuyên môn trong kỹ thuật Audio - Video
ADSR	Attack - Decay - Sustain - Release	Mô hình đường bao biên độ 4 giai đoạn theo thời gian trong thiết kế tiếng động
AES	Audio Engineering Society	Hội Kỹ sư Âm thanh Quốc tế — tổ chức ban hành các tiêu chuẩn âm thanh chuyên nghiệp
DAW	Digital Audio Workstation	Trạm làm việc âm thanh số (Pro Tools, DaVinci Resolve Fairlight, Nuendo, Reaper)
dB	Decibel	Đơn vị logarit biểu diễn tỷ số giữa hai đại lượng biên độ hoặc công suất
dB SPL	Decibel Sound Pressure Level	Mức áp suất âm thanh trong không khí quy chiếu theo ngưỡng nghe 20 µPa
dBFS	Decibels Relative to Full Scale	Thang đo decibel số toàn dải; 0 dBFS là trần tuyệt đối của hệ thống kỹ thuật số
dBTP	Decibels True Peak	Mức đỉnh thực của tín hiệu số có tính toán hiện tượng liên mẫu (inter-sample peaks)
dBA / dBC	A-weighted / C-weighted Decibels	Mức đo decibel qua mạch lọc bù trừ đặc tuyến độ nhạy của tai người
EBU	European Broadcasting Union	Hiệp hội Phát thanh Truyền hình Châu Âu (tác giả tiêu chuẩn EBU R128)
HRTF	Head-Related Transfer Function	Hàm truyền liên quan đến đầu người — cơ sở định vị không gian 3 chiều và âm thanh nhị nhĩ
ILD	Interaural Level Difference	Độ chênh lệch mức áp suất âm thanh giữa hai tai (chủ đạo ở tần số cao > 1,5 kHz)
ITD	Interaural Time Difference	Độ chênh lệch thời gian âm thanh truyền đến hai tai (chủ đạo ở tần số thấp < 1,5 kHz)
ITU-R	ITU Radiocommunication Sector	Phân ban Vô tuyến của Liên minh Viễn thông Quốc tế (tác giả chuẩn BS.1770)
LKFS / LUFS	Loudness K-weighted / Units Full Scale	Đơn vị đo lường độ to cảm nhận chuẩn hóa theo ITU-R BS.1770
LRA	Loudness Range	Dải độ to — chỉ số mô tả độ động giữa đoạn to và đoạn nhỏ của chương trình (đơn vị LU)
NC / NR	Noise Criteria / Noise Rating	Đường cong tiêu chí đánh giá mức độ ồn nền của phòng thu

Ký hiệu viết tắt	Thuật ngữ tiếng Anh đầy đủ	Ý nghĩa chuyên môn trong kỹ thuật Audio - Video
NIOSH	National Institute for Occupational Safety and Health	Viện Quốc gia về An toàn và Sức khỏe Nghề nghiệp Hoa Kỳ
RMS	Root Mean Square	Giá trị căn bậc hai trung bình bình phương (năng lượng hiệu dụng của tín hiệu)
RT60	Reverberation Time 60 dB	Thời gian để năng lượng âm vang suy giảm 60 dB sau khi ngắt nguồn âm
SMPTE	Society of Motion Picture and Television Engineers	Hiệp hội Kỹ sư Điện ảnh và Truyền hình Quốc tế
SNR	Signal-to-Noise Ratio	Tỷ số tín hiệu trên tạp âm (đo bằng dB)

CHƯƠNG 12
TRÍ TUỆ NHÂN TẠO VÀ CÔNG NGHỆ MỚI

Khung đánh giá · Từng khâu dây chuyền · Sản xuất ảo · Pháp lý

Đây là chương cuối của cuốn sách và là chương duy nhất mang tính chất sớm lạc hậu. Mười một chương trước bàn về những đối tượng thay đổi chậm như sóng âm, định lý lấy mẫu, mắt người và các hàm truyền, còn chương này bàn về những công cụ đang thay đổi ngay trong lúc người đọc cầm sách.

Trong tình huống ấy, cách làm trung thực nhất của một cuốn sách là nêu rõ đặc điểm này, đồng thời chọn vật liệu biên soạn sao cho phần thân của chương tồn tại lâu hơn các công cụ được nhắc tới. Chương thực hiện yêu cầu đó bằng cách được trình bày cụ thể ở mục 12.1.

Chương 10 trình bày hậu kỳ ở mức quy trình và hệ thống. Chương 11 trình bày công nghệ hiển thị, trong đó ba mục cuối bàn về tường LED như một thiết bị. Chương 12 nối tiếp cả hai chương: chương xem xét các công cụ mới xen vào từng khâu của dây chuyền sản xuất, phần mềm chạy trên tường LED và những nghĩa vụ pháp lý mới đi kèm.

Mục tiêu của chương

Sau khi đọc xong chương này, người đọc có thể:

1. Phân biệt ba loại tác vụ (nhận dạng, biến đổi, tạo sinh) và ba mức can thiệp (trợ giúp, bán tự động, tự động) của máy vào một khâu công việc.
2. Xác định vị trí AI xen vào chín trạm của dây chuyền và ba mốc bàn giao không thể giao cho máy.
3. Tính tỉ lệ kết quả dùng được, thời gian thực, dung lượng kho và chi phí thực của một công cụ tạo sinh theo các công thức (12.1) tới (12.6).
4. Đánh giá một công cụ AI bằng sáu câu hỏi và quyết định sử dụng dựa trên điểm hoà vốn.
5. Nêu vai trò và giới hạn của AI ở tiền kỳ, trường quay, hậu kỳ hình và hậu kỳ tiếng.
6. Mô tả quy trình sản xuất ảo và ICVFX, ba vai trò mới của ê-kíp và cách đánh giá một phương tiện mới như XR.
7. Tính ngân sách độ trễ của dựng từ xa theo (12.7) và mô tả hướng nén mới dùng mạng thần kinh cùng hiệu quả của hướng này theo (12.8).
8. Nêu nghĩa vụ minh bạch và gắn nhãn nội dung do AI tạo, xuất xứ nội dung, quyền hình ảnh và giọng nói, quyền tác giả; lập hồ sơ pháp lý cho một sản phẩm có phần do AI tạo.

Cấu trúc của chương

Chương gồm 28 mục chia thành bốn phần, mỗi phần kết thúc bằng một bài thực hành. Cách chia này theo đúng khuôn của các chương trước.

Bảng 12.1 Bốn phần của chương

Phần	Mục	Nội dung	Bài thực hành
A	12.1 – 12.7	Nguyên lý và khung đánh giá: ba loại tác vụ, ba mức can thiệp, tỉ lệ kết quả dùng được, sáu câu hỏi đánh giá một công cụ	Số 1 — đo tỉ lệ kết quả dùng được
B	12.8 – 12.14	AI ở từng khâu của dây chuyền: tiền kỳ, trường quay, hậu kỳ hình, hậu kỳ tiếng — và phép tính điểm hoà vốn	Số 2 — làm một khâu bằng hai đường
C	12.15 – 12.20	Sản xuất ảo, phương tiện mới và hạ tầng: dựng cảnh thời gian thực, ICVFX, XR, IP và đám mây, nén dùng mạng thần kinh	Số 3 — kịch bản kỹ thuật cho một cảnh sản xuất ảo
D	12.21 – 12.28	Pháp lý, xuất xứ nội dung và đạo đức nghề: nghĩa vụ minh bạch, gắn nhãn, bản sao số, quyền tác giả, cách đọc dự báo	Số 4 — hồ sơ pháp lý cho một sản phẩm có phần do AI tạo

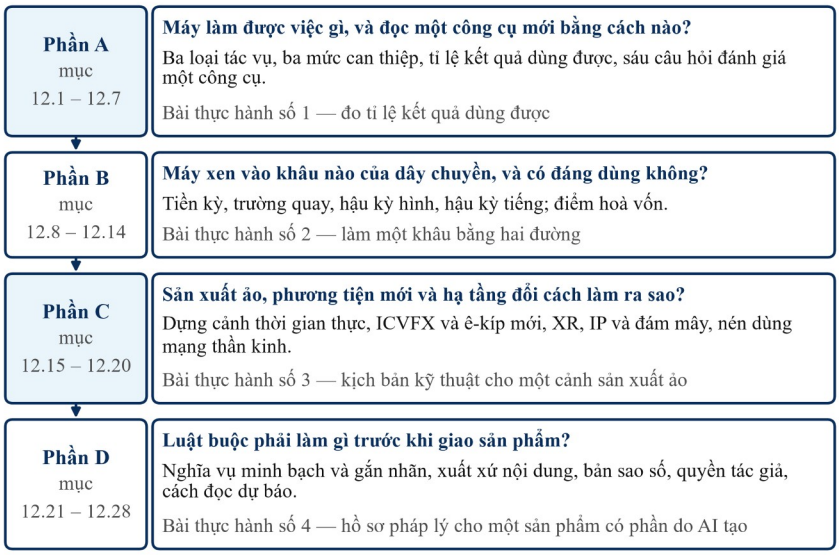
Thứ tự đọc có thể thay đổi: có thể đọc Phần D ngay sau Phần A, vì nghĩa vụ pháp lý giúp hiểu rõ hơn lý do phải phân loại tác vụ ngay từ đầu. Cách đánh số vẫn giữ A – B – C – D để thuận tiện tra cứu và việc thay đổi thứ tự đọc không ảnh hưởng đến nội dung.

12.1 Vị trí của chương trong cuốn sách

12.1.1 Chương cuối: những điều đang thay đổi

Mười một chương trước của cuốn sách bàn về những đối tượng thay đổi chậm. Sóng âm ngày nay vẫn là sóng âm của năm 1950, định lý lấy mẫu không lỗi thời, và mắt người vẫn cảm nhận màu bằng đúng ba loại tế bào cảm thụ như cách đây một triệu năm. Nhờ vậy, nội dung các chương ấy có thể sử dụng trong thời gian dài.

Chương 12 có tính chất khác hẳn vì nội dung của chương đang thay đổi ngay trong lúc người đọc cầm sách. Cách làm trung thực nhất của một cuốn sách là nêu rõ điều đó ngay từ đầu, thay vì coi như mình nằm ngoài sự biến đổi của thời gian.



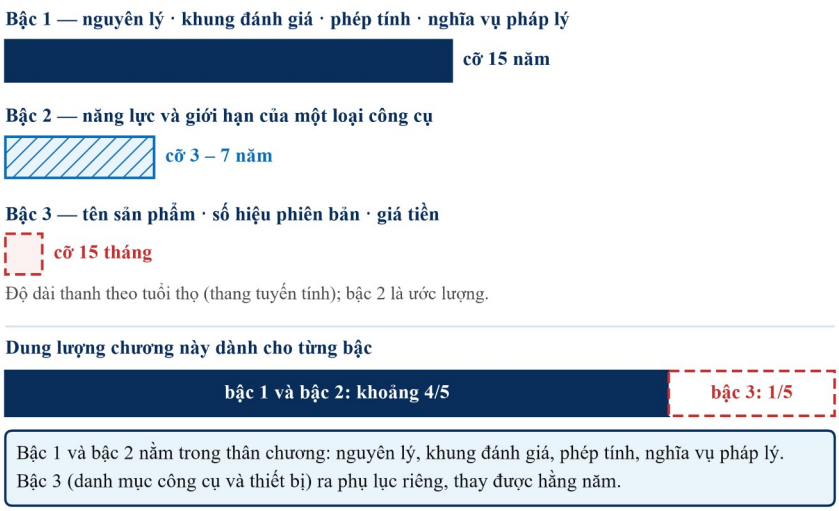
Hình 12.1 Bản đồ Chương 12 — bốn phần và câu hỏi trung tâm của mỗi phần

12.1.2 Ba bậc tuổi thọ của vật liệu viết

Một tài liệu kỹ thuật về công nghệ đang thay đổi thường lỗi thời nhanh, và nguyên nhân thường không nằm ở sự cẩu thả của người viết. Người đọc nên hiểu lý do này, vì chính họ cũng có lúc phải biên soạn những tài liệu tương tự.

Tài liệu loại này thường được xây gần như hoàn toàn bằng tên sản phẩm, số hiệu phiên bản và giá tiền — thứ vật liệu có tuổi thọ ngắn nhất trong mọi vật liệu mà một người viết sách có thể chọn. Một bảng liệt kê tám công cụ sinh video kèm tính năng và giá, sau chừng mười tám tháng, thường không còn dòng nào nguyên vẹn: sản phẩm đã lên phiên bản mới với khả năng khác hẳn, hoặc đã bị thay bằng sản phẩm của hãng khác. Một bảng thiết bị thực tế mở rộng còn dễ lỗi thời hơn: nó có thể liệt kê một sản phẩm mà nhà sản xuất đã ngừng làm.

> **QUY TẮC 12.1 — CÁCH KIỂM MỘT TÀI LIỆU KỸ THUẬT XEM NÓ SẼ SỐNG ĐƯỢC BAO LÂU**
>
> **Đọc lướt và đếm:** bao nhiêu phần trăm số câu chứa một tên sản phẩm, một số hiệu phiên bản hoặc một giá tiền?
>
> Tỉ lệ ấy là phần sẽ lỗi thời trước tiên, và nhanh hơn người biên soạn tưởng.
>
> Phép kiểm này dùng được cho mọi tài liệu, kể cả tài liệu do chính mình viết. Nó nên được làm trước khi viết, ở lúc chọn vật liệu, chứ không phải sau khi đã viết xong.



Hình 12.2 Ba bậc tuổi thọ của vật liệu viết, và tỉ lệ mà chương này dành cho từng bậc

12.1.3 Chương này được xây bằng vật liệu nào

Toàn bộ chương được chia theo ba bậc tuổi thọ ở hình trên. Khoảng bốn phần năm dung lượng dành cho vật liệu bậc một và bậc hai: nguyên lý làm việc, khung câu hỏi để đánh giá một công cụ, phép tính để quyết định, và nghĩa vụ pháp lý. Phần còn lại — danh mục công cụ và thiết bị đang có — được đẩy ra một phụ lục riêng, không đánh số cùng chương, để có thể thay hằng năm mà không phải sửa bất kỳ dòng nào của chương.

Cách chia ấy có một cái giá: đọc chương này chưa thể biết ngay nên mua phần mềm nào vào thời điểm hiện tại. Đổi lại, người đọc nắm được phương pháp tự trả lời câu hỏi ấy, kể cả với những phần mềm chưa ra đời.

> **GHI NHỚ 12.1 — BÀI HỌC QUAN TRỌNG HƠN CẢ NỘI DUNG CHƯƠNG**
>
> Một tài liệu hỏng nhanh thường không hỏng vì người viết sai, mà vì người viết chọn nhầm vật liệu.
>
> Cùng một nội dung có thể được viết bằng vật liệu tồn tại mười lăm năm (nguyên lý, khung đánh giá, phép tính) hoặc bằng vật liệu chỉ tồn tại mười lăm tháng (tên sản phẩm, số hiệu phiên bản, giá tiền).
>
> Chương này dành khoảng bốn phần năm dung lượng cho hai loại vật liệu đầu và chuyển toàn bộ loại thứ ba sang một phụ lục có thể thay hằng năm.

QUY TẮC 12.2 — CÁCH ĐỌC CÁC CON SỐ TRONG CHƯƠNG

Các mốc pháp lý (số hiệu văn bản, ngày thông qua, ngày hiệu lực, số hiệu điều khoản) được nêu đúng như văn bản quy định và phản ánh hiện trạng tới thời điểm biên soạn.

Các con số thị trường — sản lượng, doanh thu, tỉ lệ — là số liệu công bố dùng để minh hoạ một xu hướng, không phải để ghi nhớ. Điều cần nhớ là hình dáng của xu hướng và lý do của nó.

Các con số trong ví dụ tính toán (đơn giá giờ công, chi phí thiết lập, tỉ lệ kết quả dùng được) là giả thiết dùng để minh hoạ cách tính. Người đọc cần thay số liệu của mình vào, và đó là mục đích của cách trình bày này.

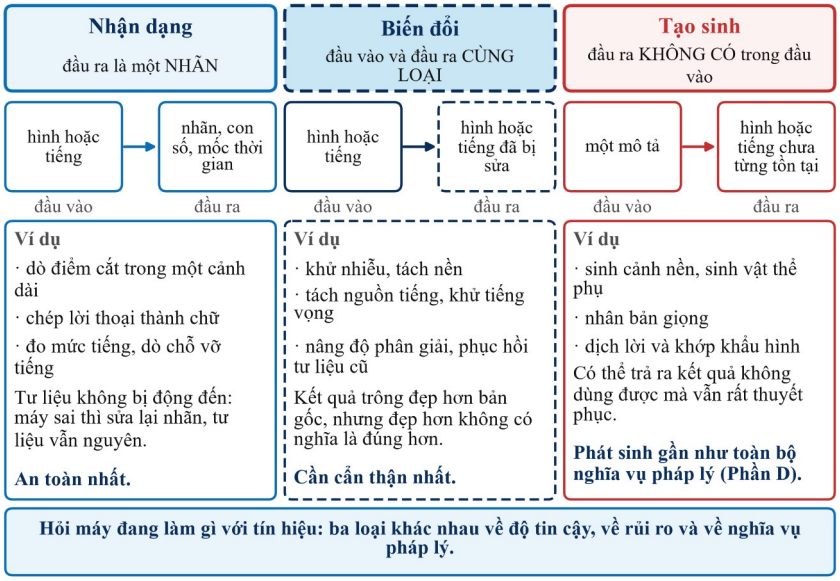
Tám công thức của chương đều được cài đặt trong một mô-đun tính toán kèm bộ tự kiểm, và mọi con số in trong chương đều được sinh ra từ mô-đun ấy.

12.2 Ba loại tác vụ mà máy làm được

12.2.1 Vì sao phải phân loại trước khi nói gì khác

Thuật ngữ "trí tuệ nhân tạo" gộp vào một chỗ những việc rất khác nhau về bản chất, về mức độ tin cậy và về hệ quả pháp lý. Một bộ lọc nhiễu và một máy sinh video đều được gọi bằng thuật ngữ ấy, nhưng bộ lọc nhiễu không phát sinh nghĩa vụ pháp lý nào, còn máy sinh video thì có. Việc trộn lẫn hai loại này là nguồn của gần như mọi hiểu lầm trong nghề.

Cách phân loại dùng suốt chương này rất đơn giản: hỏi máy đang làm gì với tín hiệu.



Hình 12.3 Ba loại tác vụ: nhận dạng, biến đổi và tạo sinh — phân theo quan hệ giữa đầu vào và đầu ra

12.2.2 Nhận dạng — đầu ra là một NHÃN

Máy nhìn vào tín hiệu và nói ra một điều gì đó về tín hiệu ấy, chứ không sửa nó. Đầu ra không phải hình hay tiếng, mà là một nhãn, một con số, một mốc thời gian.

- Dò điểm cắt trong một cảnh dài và nhận mặt để gom cảnh theo nhân vật.
- Chép lời thoại thành chữ và tách lời người nói theo từng người.
- Đo mức tiếng, dò chỗ vỡ tiếng và dò khung hình hỏng trong tệp gốc.

Đặc điểm quyết định: tư liệu không bị động đến. Máy sai thì người sửa lại nhãn, tư liệu vẫn nguyên. Đây là loại tác vụ an toàn nhất, và cũng là loại đã được dùng nhiều nhất trong sản xuất thật.

12.2.3 Biến đổi — đầu vào và đầu ra CÙNG LOẠI

Máy nhận vào tín hiệu và trả ra tín hiệu cùng loại đã bị sửa: hình ra hình, tiếng ra tiếng. Khử nhiễu, tách nền, tách nguồn tiếng, khử tiếng vọng, ổn định hình rung, nâng độ phân giải và phục hồi tư liệu cũ đều thuộc loại này.

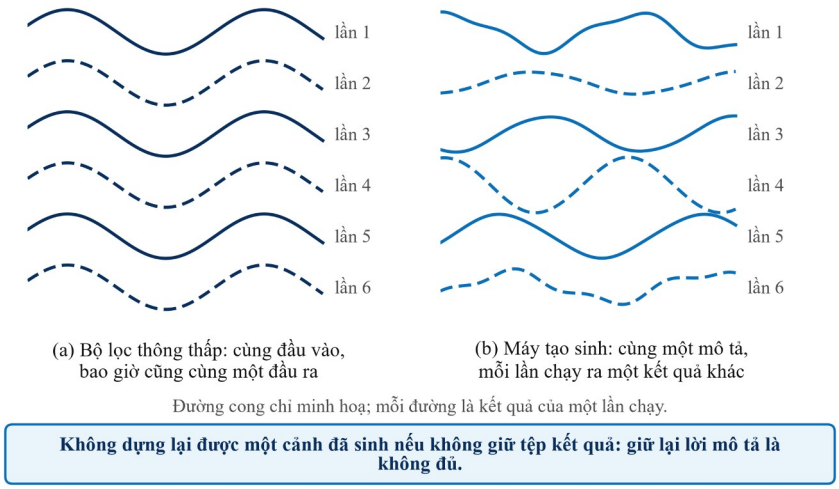
Đây là nhóm cần cẩn thận nhất, vì nó có một tính chất dễ đánh lừa: kết quả trông đẹp hơn bản gốc. Nhưng đẹp hơn không có nghĩa là đúng hơn. Khi máy nâng độ phân giải một đoạn tư liệu, nó không tìm lại được chi tiết đã mất — nó bịa ra chi tiết hợp lý theo những gì nó đã học. Với một cảnh nền hư cấu, bịa là chấp nhận được. Với một đoạn phim tài liệu lịch sử, bịa là làm sai lệch tư liệu.

> **QUY TẮC 12.3 — QUY TẮC PHÂN BIỆT, DÙNG ĐƯỢC CHO MỌI TÁC VỤ BIẾN ĐỔI**
>
> Nâng cấp tư liệu hư cấu thì được.
>
> Nâng cấp tư liệu tài liệu thì phải ghi rõ đã nâng cấp, ghi rõ bằng công cụ nào, và giữ lại bản gốc chưa xử lý.
>
> Bản đã nâng cấp không được thay thế bản gốc trong kho lưu trữ.

12.2.4 Tạo sinh — đầu ra KHÔNG CÓ trong đầu vào

Máy nhận vào một mô tả và trả ra hình hoặc tiếng chưa từng tồn tại. Sinh cảnh nền, sinh vật thể phụ, nhân bản giọng, dịch lời và khớp khẩu hình đều thuộc loại này.

Đây là nhóm phát sinh gần như toàn bộ nghĩa vụ pháp lý bàn ở Phần D, và cũng là nhóm mà chỉ số đánh giá ở mục 12.5 có vai trò quan trọng nhất, vì chỉ ở nhóm này máy mới có thể trả ra kết quả không dùng được nhưng vẫn trông rất thuyết phục.



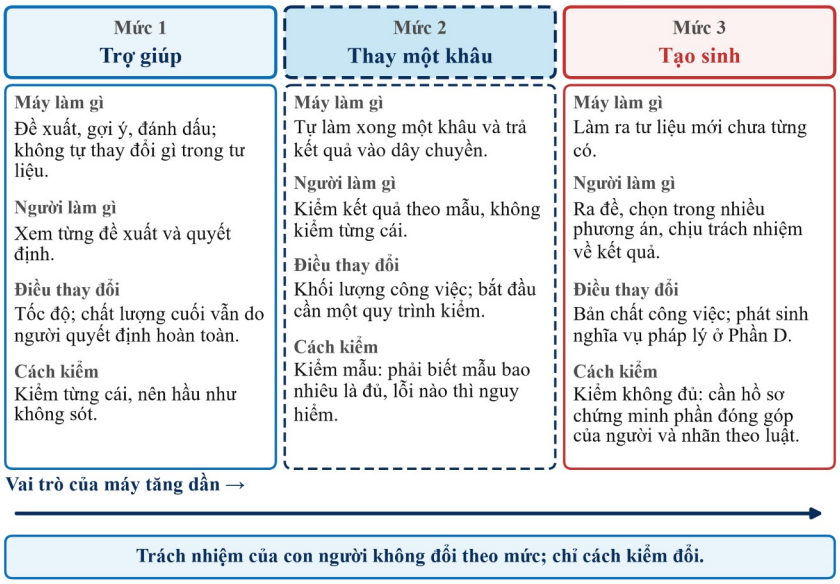
Hình 12.4 Tính không lặp lại của tác vụ tạo sinh: cùng một mô tả, mỗi lần chạy ra một kết quả khác

Tính không lặp lại ở hình trên là điểm khác biệt lớn nhất so với mọi công cụ mà mười một chương trước đã nói. Một bộ lọc thông thấp cho cùng một đầu vào thì luôn cho cùng một đầu ra. Một máy tạo sinh thì không. Hệ quả thực hành rất nặng: Không thể dựng lại một cảnh đã sinh nếu không giữ tệp kết quả — giữ lại lời mô tả là không đủ.

12.3 Ba mức can thiệp vào một khâu công việc

12.3.1 Trục thứ hai, độc lập với trục thứ nhất

Ba loại tác vụ ở mục trước nói về việc máy làm gì. Ba mức can thiệp ở mục này nói về việc máy đứng ở đâu trong quy trình — nó phụ giúp người, thay người ở một khâu, hay tự làm ra sản phẩm. Hai trục này độc lập với nhau, và cần cả hai mới định vị được một ứng dụng cụ thể.



Hình 12.5 Ba mức can thiệp: trợ giúp, thay một khâu, tạo sinh — và trách nhiệm tương ứng của con người ở mỗi mức

Bảng 12.2 Ba mức can thiệp và điều thay đổi theo từng mức

Mức	Máy làm gì	Người làm gì	Điều thay đổi
1 — Trợ giúp	Đề xuất, gợi ý, đánh dấu. Không tự thay đổi gì trong tư liệu	Xem từng đề xuất và quyết định	Tốc độ. Chất lượng cuối vẫn do người quyết định hoàn toàn
2 — Thay một khâu	Tự làm xong một khâu và trả kết quả vào dây chuyền	Kiểm kết quả theo mẫu, không kiểm từng cái	Khối lượng công việc. Bắt đầu cần một quy trình kiểm
3 — Tạo sinh	Làm ra tư liệu mới chưa từng có	Ra đề, chọn trong nhiều phương án, chịu trách nhiệm về kết quả	Bản chất công việc. Phát sinh nghĩa vụ pháp lý ở Phần D

12.3.2 Vì sao phải nhớ mức can thiệp

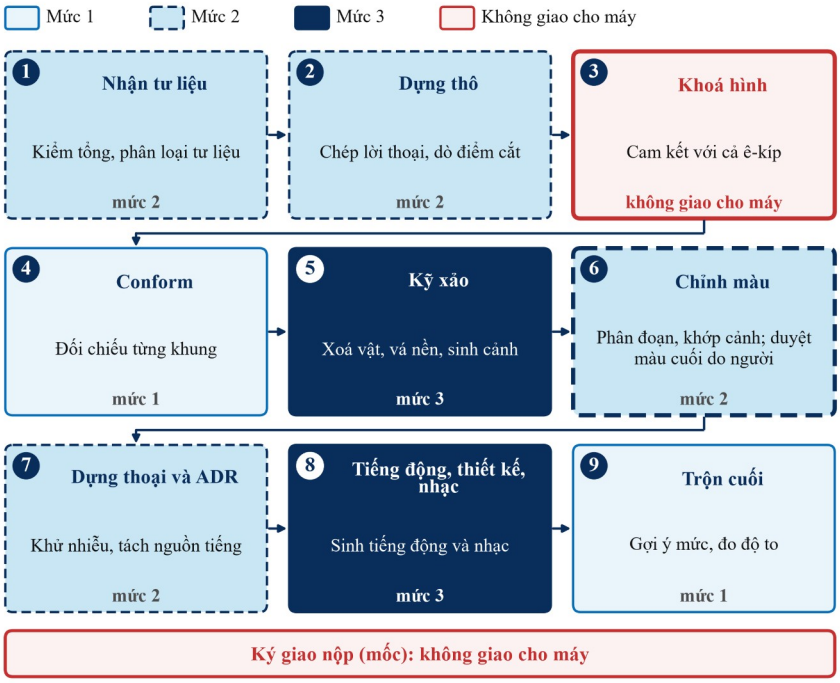
Vì trách nhiệm không thay đổi theo mức, nhưng cách kiểm thì có. Ở mức một, người kiểm từng cái nên hầu như không sót. Ở mức hai, người chỉ kiểm mẫu, nên phải biết mẫu bao nhiêu là đủ và lỗi loại nào thì nguy hiểm. Ở mức ba, kiểm không đủ nữa — phải có hồ sơ chứng minh phần đóng góp của con người, và phải có nhãn theo luật.

> **QUY TẮC 12.4 — CÂU HỎI TỰ KIỂM TRƯỚC MỖI CÔNG VIỆC**
>
> **Máy đang làm loại tác vụ nào:** nhận dạng, biến đổi hay tạo sinh?
>
> **Máy đang ở mức can thiệp nào:** trợ giúp, thay một khâu hay tạo sinh?
>
> Hai câu trả lời ấy quyết định cách kiểm tra, nội dung cần ghi chép và việc có phát sinh nghĩa vụ gắn nhãn hay không.

12.4 AI xen vào đâu trong dây chuyền

12.4.1 Chín trạm của dây chuyền, đọc lại một lần nữa

Cuốn sách đã đi qua toàn bộ dây chuyền sản xuất trong mười một chương. Mục này sắp xếp lại dây chuyền ấy thành chín trạm và đánh dấu những trạm mà máy đã xen vào, cùng mức độ xen vào.



Hình 12.6 Chín trạm của dây chuyền và mức can thiệp của máy ở từng trạm

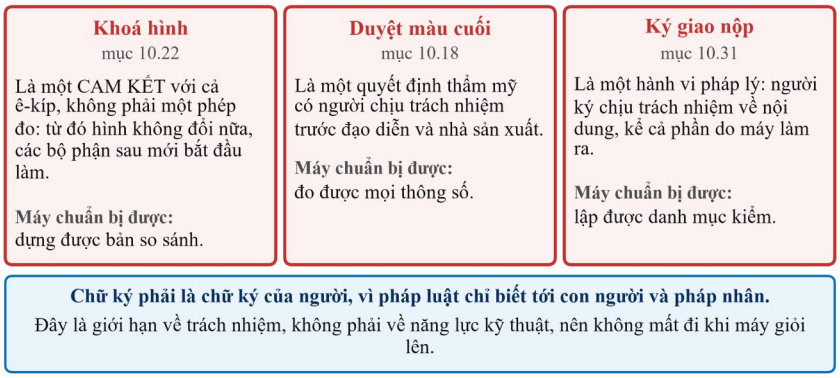
Đọc hình trên theo chiều ngang sẽ thấy một quy luật: máy xen vào mạnh nhất ở hai đầu — khâu chuẩn bị trước khi quay, và khâu xử lý sau khi quay — còn khâu giữa, tức là ngày quay thật, gần như không đổi. Lý do không phải kỹ thuật: ngày quay là chỗ tập trung con người, thiết bị và tiền, nên là chỗ ít ai dám thử nghiệm nhất.

12.4.2 Ba điều mà bản đồ này nói ra

- **Không có trạm nào bị xoá.** Máy chưa xoá bỏ trạm nào trong chín trạm, mà chỉ thay đổi cách làm bên trong từng trạm và phân bố lại khối lượng công việc giữa các trạm.
- **Khối lượng công việc được dời chỗ chứ không mất đi.** Việc giảm ở một trạm thường làm tăng việc ở trạm khác, rõ nhất là ở sản xuất ảo, bàn ở mục 12.16.
- **Chỗ máy vào sâu nhất là chỗ có nhiều phương án.** Máy mạnh ở việc sinh nhiều phương án nhanh; nó yếu ở việc chọn. Mà chọn mới là phần khó của nghề.

12.4.3 Ba mốc không giao được cho máy

Chương 10 đã nêu ba mốc bàn giao bắt buộc của một dự án: khoá hình, duyệt màu cuối và ký giao nộp. Ba mốc này vẫn giữ nguyên sau khi máy xen vào khắp dây chuyền, và lý do cần được nêu rõ.



Hình 12.7 Ba mốc bàn giao và lý do chúng không giao được cho máy

Cả ba mốc đều không phải phép đo mà là cam kết hoặc quyết định có người chịu trách nhiệm. Máy chuẩn bị được mọi thứ cho ba mốc ấy: nó lập được danh mục kiểm, dựng được bản so sánh, đo được mọi thông số. Nhưng chữ ký thì phải là chữ ký của người, vì pháp luật chỉ biết tới con người và pháp nhân. Đây là một giới hạn về trách nhiệm, không phải về năng lực kỹ thuật — nên nó sẽ không mất đi khi máy giỏi lên.

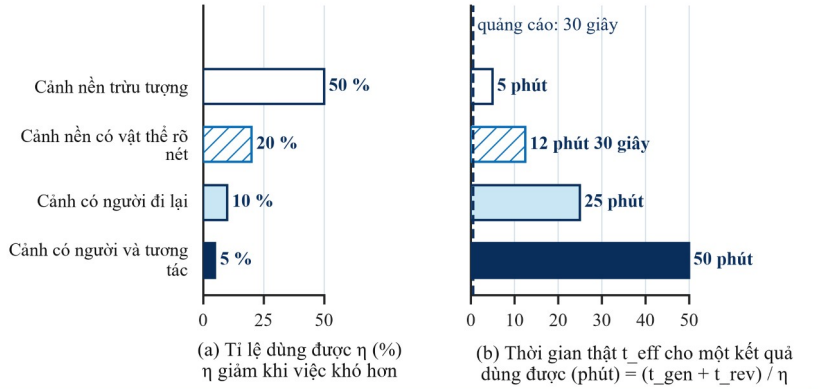
12.5 Hai số đo của một công cụ tạo sinh

12.5.1 Tỉ lệ kết quả dùng được

Mọi tài liệu quảng cáo đều nói công cụ sinh ra một cảnh trong bao lâu. Không tài liệu nào nói phải sinh mấy lần mới được một cảnh dùng được. Con số thứ hai ấy mới là con số quyết định, và nó phải tự đo lấy.

$$\eta = \frac{n_{\text{ok}}}{n_{\text{gen}}} \tag{12.1}$$

trong đó n_{ok} là số kết quả giữ lại được và n_{gen} là số lần đã sinh. Tỉ lệ η không phải thuộc tính của công cụ mà phụ thuộc vào công cụ, loại công việc, người ra đề và mức khắt khe của người duyệt. Với cùng một công cụ, một cảnh nền trừu tượng có thể cho η rất cao, trong khi một cảnh có người đi lại cho η rất thấp.



Hình 12.8 Tỉ lệ kết quả dùng được thay đổi thế nào theo độ khó của việc

12.5.2 Thời gian thật cho một kết quả dùng được

Có η rồi thì tính được con số mà người ta thật sự cần: thời gian bỏ ra cho một kết quả dùng được, chứ không phải cho một lần sinh.

$$t_{\text{eff}} = \frac{t_{\text{gen}} + t_{\text{rev}}}{\eta} \tag{12.2}$$

trong đó t_{gen} là thời gian máy sinh một lần và t_{rev} là thời gian người xem và quyết định giữ hay bỏ một lần. Phép chia cho η xuất phát từ việc trung bình phải thực hiện 1/η lần mới được một kết quả dùng được.

> **VÍ DỤ 12.1 — VÍ DỤ TÍNH — VÀ CHỖ TRỰC GIÁC ĐÁNH LỪA**
>
> Máy sinh một cảnh 8 giây mất 30 giây, người xem và quyết định mất 120 giây, và tỉ lệ dùng được đo được là η = 20 %.
>
> Thời gian thật cho một cảnh dùng được là t_{eff} = (30 + 120) / 0,2 = 750 giây, tức 12 phút 30 giây.
>
> Quảng cáo nêu "30 giây một cảnh", trong khi thời gian thực tế là 12 phút 30 giây, gấp 25 lần. Cả hai con số đều đúng; chúng chỉ trả lời hai câu hỏi khác nhau.
>
> **Chú ý thêm:** phần lớn thời gian nằm ở khâu kiểm, không ở khâu sinh. Nên muốn nhanh hơn thì phải rút ngắn khâu kiểm, không phải mua máy sinh nhanh hơn.

12.5.3 Cách tự đo η mà không tốn nhiều công

Việc đo không đòi hỏi một cuộc thử nghiệm lớn. Cách làm đủ dùng là chọn mười việc điển hình của loại công việc định thực hiện, mỗi việc cho máy sinh tới khi được một kết quả chấp nhận được hoặc tới lần thứ mười thì dừng, rồi ghi số lần sinh của từng việc. Tổng số lần sinh là n_{gen}, còn số việc thành công là n_{ok}.

Điều quan trọng hơn con số: Phải định nghĩa trước thế nào là "dùng được", bằng chữ, trước khi bắt đầu sinh. Nếu định nghĩa ấy được nới ra trong lúc làm — và nó rất hay bị nới ra — thì η đo được sẽ đẹp mà vô nghĩa.

12.6 Sáu câu hỏi để đánh giá một công cụ

12.6.1 Khung dùng được cho công cụ chưa ra đời

Sáu câu hỏi dưới đây áp dụng được cho một công cụ chưa tồn tại vào lúc viết chương này, vì chúng không hỏi về tính năng mà hỏi về chỗ đứng của công cụ trong công việc thật.



Hình 12.9 Sáu câu hỏi để đánh giá một công cụ, và thứ tự hỏi

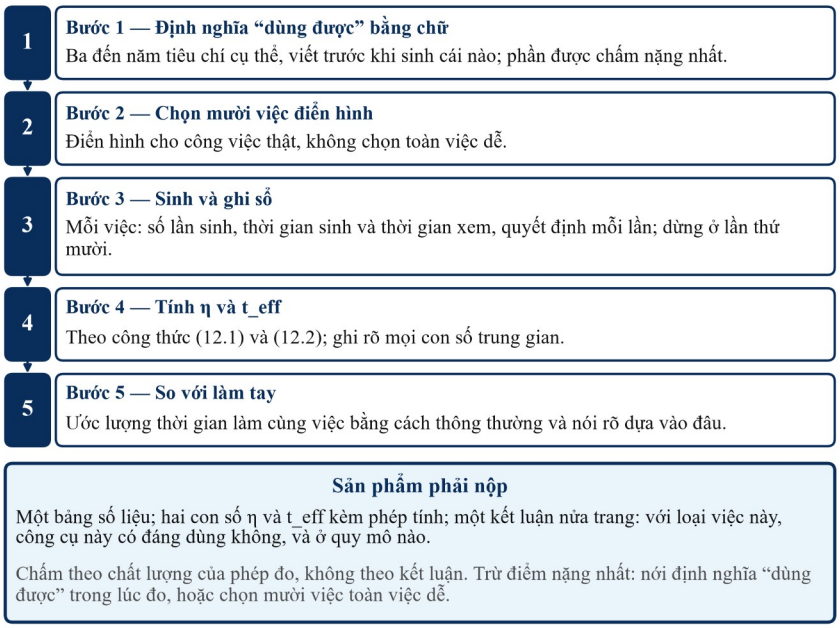
Bảng 12.3 Sáu câu hỏi, điều cần tìm và dấu hiệu đáng lo

Câu hỏi	Điều cần tìm	Dấu hiệu đáng lo
1. Nó thay khâu nào, không phải nó làm được gì	Một khâu cụ thể trong chín trạm ở mục 12.4, gọi được tên	Không chỉ ra được khâu nào — nghĩa là chưa có chỗ dùng thật
2. Tỉ lệ dùng được là bao nhiêu, đo bằng việc của mình	Một con số η tự đo trên mười việc điển hình	Chỉ có video giới thiệu, không cho thử với tư liệu của mình
3. Kết quả có lặp lại được không	Cùng một đề, cùng một hạt giống thì ra cùng kết quả	Không lặp lại được và cũng không cho giữ hạt giống
4. Tư liệu của mình đi đâu	Điều khoản nói rõ tư liệu tải lên có bị dùng để huấn luyện không, lưu bao lâu, ai xem được	Điều khoản mơ hồ, hoặc cho phép dùng tư liệu người dùng để huấn luyện mà không hỏi
5. Nó cắm vào dây chuyền hiện có bằng cách nào	Định dạng vào ra chuẩn, xuất được ra thứ mà phần mềm dựng đọc thẳng	Chỉ chạy trong trình duyệt của hãng, xuất ra định dạng riêng
6. Hỏng thì sao	Hãng đóng cửa, đổi giá, đổi điều khoản thì công việc đang làm có chết không	Toàn bộ tư liệu và quy trình bị khoá trong dịch vụ của một hãng

12.6.2 Câu hỏi số sáu, thường bị bỏ qua nhất

Trong sáu câu, câu cuối ít được hỏi nhất nhưng lại gây thiệt hại lớn nhất khi bị bỏ qua. Một dự án kéo dài hai năm mà toàn bộ tư liệu trung gian nằm trong dịch vụ của một hãng nhỏ là một dự án đang đặt cược vào sự tồn tại của hãng ấy. Cách phòng đơn giản và nên áp dụng từ ngày đầu: mọi kết quả đều được xuất ngay ra định dạng chuẩn và lưu ở kho riêng, còn dịch vụ bên ngoài chỉ được coi là máy chạy tạm.

12.7 Bài thực hành số 1 — đo tỉ lệ dùng được



Hình 12.10 Bài thực hành số 1 — cấu trúc bài và kết quả cần có

12.7.1 Đề bài

Chọn một công cụ tạo sinh đang có và một loại việc cụ thể trong nghề — chẳng hạn sinh hình phân cảnh cho một kịch bản ngắn, hoặc sinh một cảnh nền tĩnh. Đo tỉ lệ kết quả dùng được của công cụ ấy cho loại việc ấy, rồi tính thời gian thật cho một kết quả dùng được.

12.7.2 Các bước bắt buộc

- **Bước 1: định nghĩa "dùng được" bằng chữ.** Viết ra ba tới năm tiêu chí cụ thể trước khi sinh kết quả đầu tiên, và giữ bản định nghĩa này kèm kết quả, vì đây là phần quan trọng nhất của bài.
- **Bước 2: chọn mười việc điển hình.** Các việc phải điển hình cho công việc thật, không được chọn toàn việc dễ.
- **Bước 3: sinh và ghi sổ.** Với mỗi việc, ghi số lần sinh, thời gian sinh mỗi lần, thời gian xem và quyết định mỗi lần; dừng ở lần thứ mười nếu chưa đạt kết quả dùng được.
- **Bước 4: tính η và t_{eff}.** Dùng hai công thức ở mục 12.5 và ghi rõ mọi con số trung gian.

– **Bước 5: so với làm tay.** Ước lượng thời gian làm cùng việc ấy bằng cách thông thường và nêu rõ căn cứ của ước lượng đó.

12.7.3 Kết quả cần có và cách tự đánh giá

Kết quả gồm một bảng số liệu, hai con số η và t_{eff} kèm phép tính, và một kết luận dài nửa trang: với loại việc này, công cụ này có đáng dùng hay không, và ở quy mô công việc nào thì đáng.

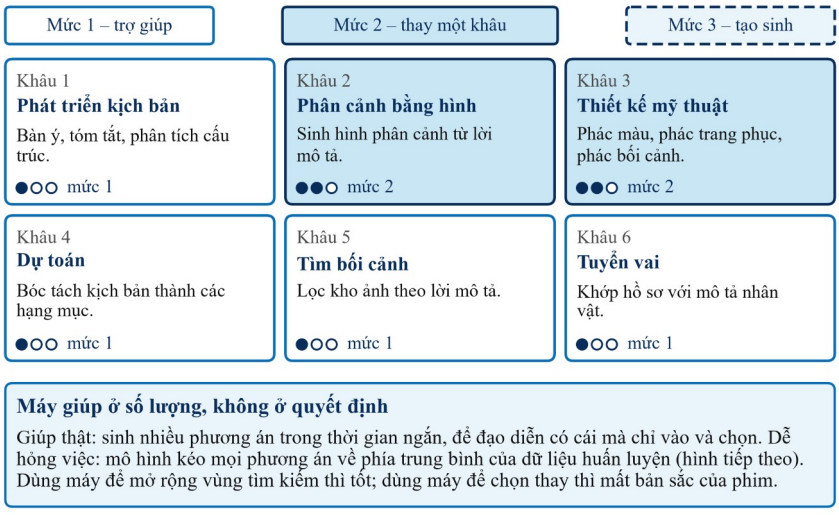
> **THỰC HÀNH 12.1 — CÁCH TỰ ĐÁNH GIÁ**
>
> Đánh giá theo chất lượng của phép đo, không theo kết luận.
>
> Một bài kết luận "không nên dùng" nhưng đo cẩn thận, định nghĩa rõ ràng và ghi sổ đầy đủ có giá trị hơn một bài kết luận "rất tốt" mà không có số liệu.
>
> Hai lỗi nặng nhất là nới định nghĩa "dùng được" trong lúc đo và chọn mười việc toàn việc dễ.

Nếu có điều kiện, bài này nên làm hai người, và nên để hai người đánh giá chéo kết quả của nhau theo cùng bộ tiêu chí. Chênh lệch giữa hai người đánh giá chính là thứ đáng bàn nhất sau khi làm xong: nó cho thấy phần chủ quan trong mọi con số η mà người ta công bố.

12.8 Tiền kỳ

12.8.1 Sáu khâu và mức can thiệp của máy

Tiền kỳ là khâu máy xen vào sớm nhất và sâu nhất, vì đây là khâu có nhiều phương án nhất và ít rủi ro nhất: một hình phân cảnh hỏng thì bỏ đi, không tốn gì ngoài thời gian.



Hình 12.11 Sáu khâu của tiền kỳ và mức can thiệp của máy ở từng khâu

12.8.2 Chỗ máy giúp thật

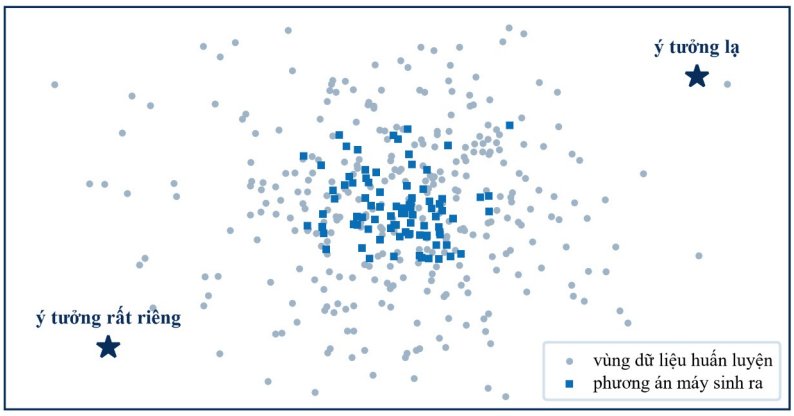
Cái mà máy làm tốt hơn người ở tiền kỳ không phải chất lượng, mà là số lượng phương án trong một đơn vị thời gian. Một hoạ sĩ phân cảnh giỏi vẽ mười khuôn hình trong một buổi. Máy sinh một trăm khuôn hình trong mười phút.

Không khuôn nào của máy bằng khuôn của hoạ sĩ, nhưng có một trăm cái để đạo diễn chỉ vào và nói "gần cái này, nhưng lùi máy ra và hạ thấp xuống".

Nói cách khác: máy không thay hoạ sĩ phân cảnh; nó thay giai đoạn mò mẫm trước khi hoạ sĩ bắt đầu vẽ. Đó là một khâu có thật, tốn thời gian có thật, và trước nay chưa ai đặt tên cho nó.

12.8.3 Chỗ máy hỏng việc — bẫy trung bình

Một mô hình tạo sinh học từ rất nhiều tác phẩm đã có, và cái nó học được là xu hướng chung của khối tư liệu ấy. Cho nên khi được yêu cầu sinh một cảnh, nó có khuynh hướng trả ra thứ nằm gần trung tâm của phân bố — bố cục quen thuộc, ánh sáng quen thuộc, màu quen thuộc.



Dữ liệu mô phỏng để minh hoạ cơ chế: máy sinh dày ở giữa, thưa ở rìa; khoảng cách trung bình tới tâm của phương án máy sinh chỉ bằng 40 % của vùng dữ liệu.

Vì sao điều này xảy ra

Mô hình trả về thứ có xác suất cao theo dữ liệu đã học; thứ có xác suất cao nhất chính là thứ đã thấy nhiều lần. Nên các phương án máy sinh ra thường giống nhau hơn là phương án của những người khác nhau.

Cách dùng đúng

Dùng máy để mở rộng vùng tìm kiếm, rồi để người đẩy các phương án ra khỏi vùng trung bình. Không dùng máy để chọn phương án cuối. Câu kiểm tra: phương án này có thể là của bất kỳ bộ phim nào không? Nếu có, nó chính là cái trung bình.

Hình 12.12 Bẫy trung bình: mô hình kéo mọi phương án về phía tâm của dữ liệu huấn luyện

CẢNH BÁO 12.1 — VÌ SAO ĐÂY LÀ VẤN ĐỀ NGHIÊM TRỌNG CHỨ KHÔNG PHẢI CHUYỆN THẨM MỸ

Một bộ phim có bản sắc là một bộ phim đi chệch khỏi trung bình ở những chỗ có chủ ý.

Máy kéo kết quả về trung bình một cách âm thầm: từng khuôn hình đều "ổn" nên không ai thấy điều gì sai, cho tới khi xem cả phim và nhận ra nó giống mọi phim khác.

Cách phòng: dùng máy để sinh phương án ở giai đoạn mò mẫm, nhưng không dùng nó ở giai đoạn quyết định hình thức của phim. Hai giai đoạn ấy phải tách bạch, và người quyết định phải biết mình đang ở giai đoạn nào.

12.9 Trường quay

12.9.1 Khâu ít gây tranh cãi nhất

Ở trường quay, gần như toàn bộ ứng dụng của máy nằm ở mức trợ giúp: nó đề xuất, nó ghi chép, nó cảnh báo, nhưng người vẫn bấm nút. Đó là lý do khâu này ít gây tranh cãi nhất trong cả dây chuyền.



Hình 12.13 Bốn ứng dụng ở trường quay, tất cả đều ở mức trợ giúp

Bảng 12.4 Bốn ứng dụng ở trường quay

Ứng dụng	Máy làm gì	Điều được và điều phải cảnh giác
Bám nét tự động theo mặt hoặc theo mắt	Nhận ra khuôn mặt trong khuôn hình và giữ nét vào đó	Được: cứu được nhiều cảnh khó ở khẩu độ mở lớn. Cảnh giác: máy chọn theo xác suất — nó bám khuôn mặt rõ nhất, không bám khuôn mặt quan trọng nhất với cảnh
Ghi siêu dữ liệu ống kính	Ghi tiêu cự, khẩu độ, khoảng nét, méo hình theo từng khung hình	Được: hậu kỳ ghép hình khớp ngay, đỡ được nhiều ngày dò lại. Đây là ứng dụng có lợi nhất và ít rủi ro nhất ở khâu này
Chép lời thoại tại chỗ	Chuyển tiếng thu tại trường quay thành chữ, gắn mốc thời gian	Được: phòng dựng tìm được câu thoại bằng cách gõ chữ. Cảnh giác: bản chép là để tìm kiếm, không phải để làm phụ đề — độ chính xác ở trường quay ồn thường không đủ
Theo dõi khớp nối cảnh	So khuôn hình hiện tại với khuôn hình đã quay để dò chỗ lệch	Được: bắt được lỗi đạo cụ đặt sai chỗ, cúc áo cài lệch. Cảnh giác: nó không hiểu chỗ nào lệch thì có hại, nên báo rất nhiều báo động giả

12.9.2 Một cảnh báo đáng nhớ về bám nét tự động

Bám nét tự động là ứng dụng dễ tin nhất và cũng là ứng dụng đã làm hỏng nhiều cảnh nhất. Vấn đề nằm ở chỗ máy tối ưu một thứ khác với thứ mà đạo diễn muốn: nó giữ nét vào khuôn mặt nào rõ nhất và ổn định nhất, trong khi đạo diễn có thể đang muốn nét nằm ở người đứng phía sau, hoặc muốn nét trôi khỏi mặt để chuyển sự chú ý.

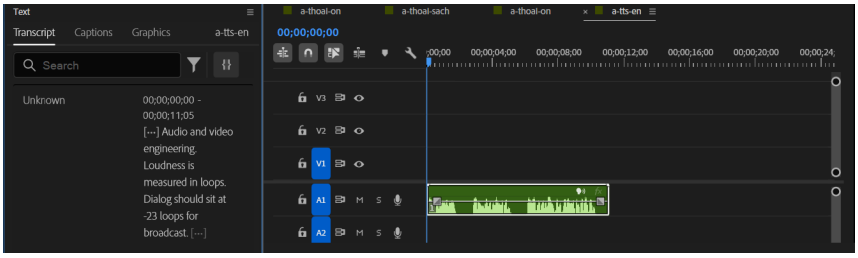
Quy tắc thực hành là dùng bám nét tự động cho các cảnh mà ý đồ trùng với tiêu chí máy tối ưu, như phỏng vấn, cảnh một người và cảnh theo dõi, và tắt nó ở mọi cảnh có chuyển nét mang nghĩa.

12.10 Hậu kỳ hình — nhận dạng và biến đổi

12.10.1 Nhóm chín muồi nhất

Đây là nhóm ứng dụng đã đi sâu nhất vào sản xuất thật, nhờ hai tính chất thuận lợi: kết quả kiểm được ngay bằng mắt, và tư liệu gốc vẫn còn nguyên nên khi hỏng có thể làm lại.

Hình 12.14 là kết quả phiên âm tự động của Premiere Pro trong bảng Text: đoạn thoại tiếng Anh được chuyển thành văn bản kèm mốc thời gian, nhưng vẫn có chỗ nghe sai (ví dụ “LUFS” thành “loops”), nhắc rằng kết quả cần được người kiểm lại.



Hình 12.14 Phiên âm tự động trong bảng Text của Adobe Premiere Pro, từ một đoạn thoại tiếng Anh do giọng tổng hợp đọc; có lỗi nhận dạng cần kiểm lại

Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Premiere Pro 2026), 07/11/2026.



Hình 12.15 Sáu ứng dụng hậu kỳ hình thuộc nhóm nhận dạng và biến đổi

Bảng 12.5 Sáu ứng dụng và mức chín muồi tới thời điểm biên soạn

Ứng dụng	Thay việc gì	Mức chín muồi
Phân đoạn tự động	Thay vòng khoá vẽ tay từng khung hình	Rất cao với người và vật thể rõ viền; còn yếu với tóc bay, khói, lưới, kính
Khử nhiễu	Thay bộ lọc khử nhiễu theo thời gian	Rất cao. Giữ được chi tiết tốt hơn hẳn bộ lọc cũ ở cùng mức khử
Nội suy khung hình cho quay chậm	Thay việc quay ở tốc độ cao ngay từ đầu	Cao với chuyển động đều; hỏng ở chỗ có vật thể che nhau
Nâng độ phân giải	Thay việc quét lại phim gốc ở độ phân giải cao hơn	Cao về mặt nhìn, nhưng xem mục 12.2.3 về ranh giới tư liệu
Phục hồi tư liệu cũ	Thay công đoạn vá xước, khử rung, khử chớp sáng làm tay	Cao, và là ứng dụng có giá trị xã hội rõ nhất của cả nhóm
Đóng khung lại theo tỉ lệ khác	Thay việc dịch khuôn hình bằng tay cho bản dọc, bản vuông	Cao cho nội dung mạng; không dùng được cho phim chiếu rạp

12.10.2 Nối lại với Chương 10

Mục 10.18 và 10.19 của Chương 10 đã nói về vòng khoá và về khử nhiễu như những công đoạn tốn công nhất của hậu kỳ hình. Nhóm ứng dụng ở đây không đổi bản chất hai công đoạn ấy — nó rút thời gian xuống. Điều đáng lưu ý cho người quản lý sản xuất: khi một công đoạn rút từ ba ngày xuống ba giờ, cái thay đổi không chỉ là chi phí mà là số lần dám sửa. Một đạo diễn biết việc sửa mất ba ngày thì không dám đổi ý; biết mất ba giờ thì đổi ý nhiều lần. Đây là chỗ mà lợi ích kỹ thuật có thể biến thành phình chi phí, nếu không ai đặt giới hạn.

12.11 Hậu kỳ hình — tạo sinh

12.11.1 Bốn ứng dụng và giới hạn thật của chúng

Đây là nhóm được nói tới nhiều nhất và cũng là nhóm bị hiểu sai nhiều nhất. Cần tách bạch hai điều: chất lượng của một cảnh ngắn, và khả năng giữ được sự nhất quán qua nhiều cảnh.

Hình 12.16 minh hoạ trạng thái trước khi tạo sinh: vùng chọn trên ảnh phong cảnh và ô nhập mô tả của Generative Fill, chưa bấm Generate.



Hình 12.16 Generative Fill trong Photoshop trước khi tạo sinh: vùng chọn và ô nhập mô tả (tuỳ chọn). Ảnh nền là ảnh cầu vồng trên núi từ Wikimedia Commons
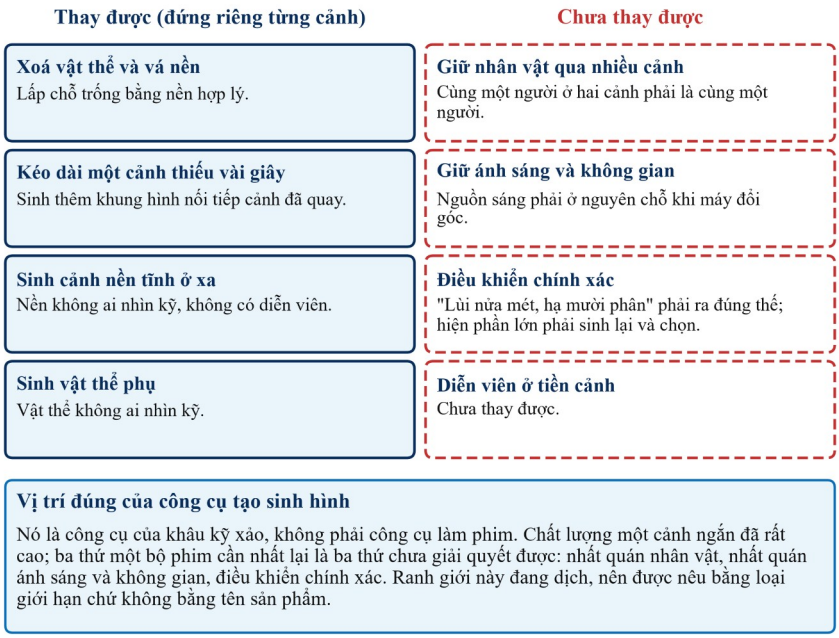
Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Photoshop 2025), 07/11/2026; ảnh nền: Muttnick, CC BY-SA 4.0, Wikimedia Commons.

Hình 12.17 là một ví dụ ảnh do mô hình tạo sinh tạo ra từ câu lệnh văn bản, dùng để quan sát chất lượng của một khung hình đơn lẻ.



Hình 12.17 Ảnh do mô hình Stable Diffusion 3.5 tạo ra từ câu lệnh văn bản (tác giả gọi là synthograph); không phải ảnh chụp

Nguồn: VulcanSphere, Public domain, Wikimedia Commons (truy cập 05/11/2026).



Hình 12.18 Bốn ứng dụng tạo sinh ở hậu kỳ hình và giới hạn của từng ứng dụng

Tới thời điểm biên soạn, chất lượng của một cảnh ngắn đứng riêng đã rất cao: một cảnh dài vài giây, độ phân giải đủ chiếu, có tiếng đi kèm đã là điều bình thường. Tuy nhiên, ba yêu cầu sau vẫn chưa được giải quyết, và đó lại chính là ba yêu cầu mà một bộ phim cần nhất:

– **Giữ nhân vật qua nhiều cảnh.** Cùng một người được quay ở hai cảnh khác nhau phải vẫn là cùng một người; đây là chỗ hỏng rõ nhất.
– **Giữ ánh sáng và không gian.** Nguồn sáng phải ở nguyên vị trí khi máy quay đổi góc, trong khi máy tạo sinh chưa có mô hình không gian mà chỉ có thói quen thống kê.
– **Điều khiển chính xác.** Khi đạo diễn yêu cầu "lùi máy ra nửa mét và hạ xuống mười phân", kết quả phải đúng như vậy, nhưng hiện nay phần lớn trường hợp phải sinh lại và chọn.

12.11.2 Kết luận thực tế về vị trí của nhóm này

> **GHI NHỚ 12.2 — VỊ TRÍ ĐÚNG CỦA CÔNG CỤ TẠO SINH HÌNH, TÍNH TỚI THỜI ĐIỂM BIÊN SOẠN**
>
> Nó là công cụ của khâu kỹ xảo, không phải công cụ làm phim.
>
> Nó thay được các việc như xoá vật thể và vá nền, kéo dài một cảnh thiếu vài giây, sinh cảnh nền tĩnh ở xa và sinh vật thể phụ không ai nhìn kỹ.
>
> Nó chưa thay được những thứ cần giữ nhất quán qua nhiều cảnh và những thứ có diễn viên ở tiền cảnh.
>
> Ranh giới này đang dịch, và đó chính là lý do mục này được viết bằng các loại giới hạn chứ không bằng tên sản phẩm.

12.11.3 Hai phép tính phải làm trước khi hứa với đoàn

Trước khi nhận làm một cảnh bằng công cụ tạo sinh, cần tính ra hai con số, vì cả hai đều dễ bị bỏ qua do không xuất hiện trong phần giới thiệu của công cụ.

Thứ nhất là thời gian sinh, tính theo số khung hình:

$$N_{\mathrm{fr}} = t \cdot f_{\mathrm{fr}} \quad ; \quad t_{\mathrm{gen}} = \frac{N_{\mathrm{fr}}}{R_{\mathrm{gen}}} \tag{12.3}$$

trong đó N_{fr} là số khung hình của cảnh, f_{fr} là tần số hình, và R_{gen} là số khung hình máy sinh được trong một giây.

Ví dụ: cảnh 8 giây ở 24 hình một giây có $N_{fr} = 192$ khung. Máy sinh được 2 khung một giây thì $t_{gen} = 96$ giây cho một lần sinh — tức chỉ bằng 0,083 lần thời gian thực. Nhân với $1/\eta$ ở mục 12.5 mới ra thời gian thật cho một cảnh dùng được.

Con số thứ hai là dung lượng kho. Sinh nhiều phương án thì kho phình ra rất nhanh, vì mỗi cảnh dùng được đi kèm nhiều cảnh bị loại:

$$S_{\mathrm{peak}} = n_{\mathrm{gen}} \cdot S_1 \quad ; \quad S_{\mathrm{keep}} = S_{\mathrm{peak}} \cdot k_{\mathrm{keep}} \tag{12.4}$$

trong đó S_1 là dung lượng một phương án và k_{keep} là tỉ lệ giữ lại sau khi lọc. Ví dụ, một cảnh 8 giây ở 50 Mbit/s nặng 0,05 GB. Sinh 10 phương án cho mỗi cảnh, cả phim có 200 cảnh, thì lúc cao điểm kho cần khoảng 100 GB; giữ lại 20 % sau khi lọc thì còn khoảng 20 GB.

QUY TẮC 12.5 — CON SỐ PHẢI CHUẨN BỊ CHỖ LÀ CON SỐ CAO ĐIỂM

Người ta quen ước lượng theo dung lượng sản phẩm cuối, tức con số sau khi lọc, nhưng kho phải chứa được con số lúc cao điểm, và hai con số ấy chênh nhau đúng $1/k_{keep}$ lần.

Kèm theo là một kỷ luật vận hành: phải có lịch dọn. Không có lịch dọn thì k_{keep} bằng 1, vì không ai dám xoá thứ mình đã bỏ công sinh ra.



(a) Bản gốc, độ phân giải thấp

(b) Nội suy thường: mờ nhưng trung thực

(c) Nâng cấp bằng mô hình: sắc nét hơn, chi tiết được sinh thêm

Quy tắc phân biệt

Bản bên phải sắc nét hơn, nhưng những chi tiết sắc nét ấy không có trong bản gốc: máy sinh ra chi tiết hợp lý theo những gì nó đã học. Nâng cấp tư liệu hư cấu thì được. Nâng cấp tư liệu tài liệu thì phải ghi rõ đã nâng cấp, bằng công cụ nào, và giữ bản gốc chưa xử lý; không để bản nâng cấp thay thế bản gốc trong kho lưu trữ.

Hình 12.19 Nâng độ phân giải: chi tiết ở bản nâng cấp là chi tiết được bịa ra, không phải chi tiết được tìm lại

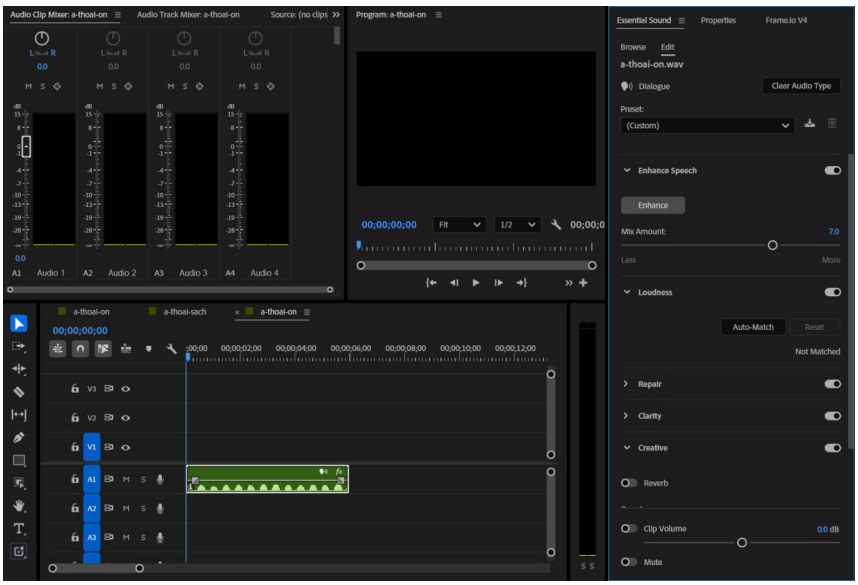
Hình trên nhắc lại một lần nữa quy tắc đã nêu ở mục 12.2.3, vì nó là quy tắc dễ quên nhất khi kết quả trông đẹp: bản bên phải sắc nét hơn, nhưng những chi tiết sắc nét ấy không có trong bản gốc. Với một cảnh hư cấu thì không sao. Với một tư liệu lịch sử thì đó là làm sai lệch tư liệu, dù không ai cố ý.

12.12 Hậu kỳ tiếng — dọn tiếng và tách nguồn

12.12.1 Nhóm mà máy đã vượt hẳn công cụ cũ

Nếu phải chỉ ra một lĩnh vực mà công cụ dùng mạng thần kinh đã vượt hẳn công cụ truyền thống thì đó là dọn tiếng. Lý do nằm ở bản chất của tín hiệu: tiếng nói có cấu trúc rất chặt và ổn định, nên một mô hình học được cấu trúc ấy có thể tách lời khỏi ồn theo cách mà không bộ lọc thiết kế thủ công nào làm được.

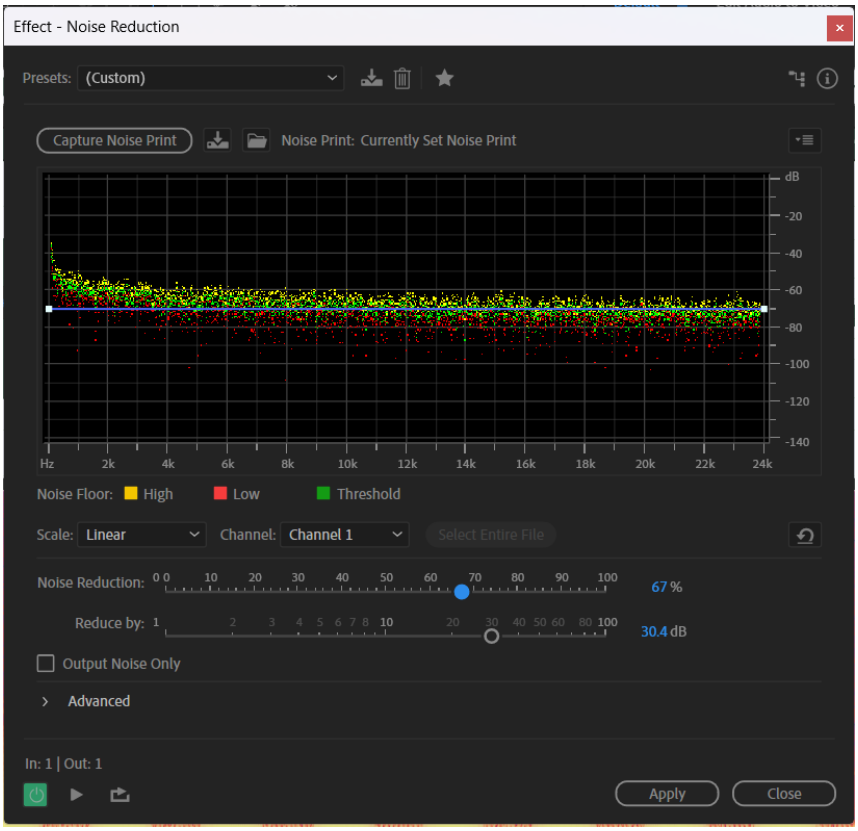
Hình 12.20 cho thấy mục Enhance Speech của bảng Essential Sound, nơi một nút bấm đủ để máy dọn tạp âm và tiếng vang của đoạn thoại, với thanh Mix Amount điều chỉnh mức can thiệp.



Hình 12.20 Enhance Speech trong Essential Sound của Adobe Premiere Pro: dọn tiếng nói bằng một nút bấm (tệp thoại mẫu tổng hợp)

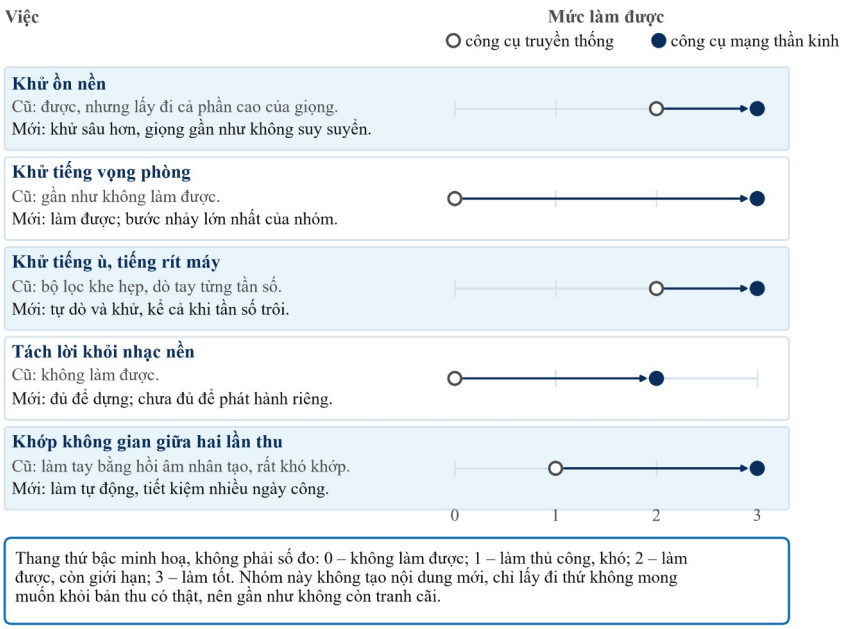
Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Premiere Pro 2026), 07/11/2026.

Để đối chiếu với công cụ mới, hình 12.19 cho thấy công cụ truyền thống tiêu biểu: hộp thoại Noise Reduction của Adobe Audition, trong đó dấu vân tay tạp âm được lấy từ một đoạn lặng ngắn rồi trừ khỏi toàn bộ bản thu.



Hình 12.21 Hộp thoại Noise Reduction (giảm tạp âm theo dấu vân tay) của Adobe Audition, áp dụng cho tệp thoại có tạp âm mẫu tổng hợp

Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Audition 2026, bản 26.5), 07/11/2026.



Hình 12.22 Năm việc dọn tiếng và mức cải thiện so với công cụ truyền thống

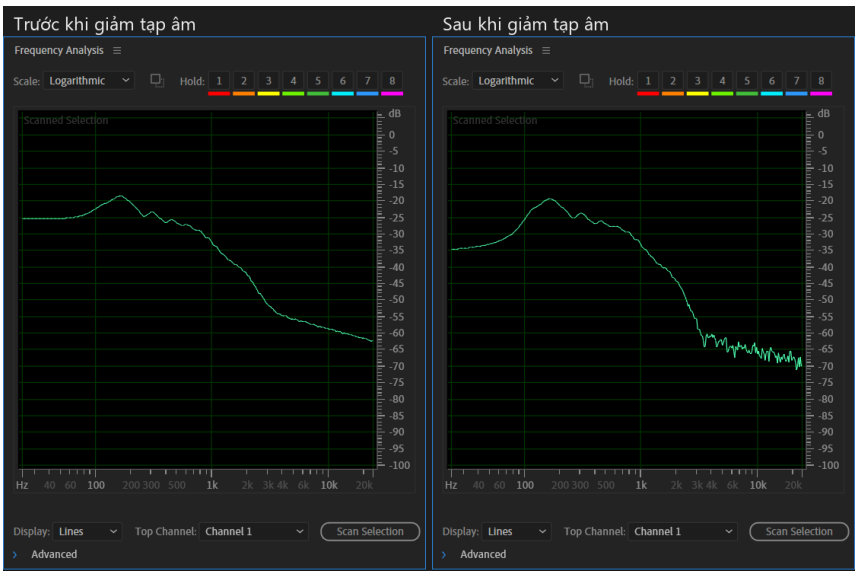
Bảng 12.6 Năm việc dọn tiếng

Việc	Công cụ cũ làm được tới đâu	Công cụ mới
Khử ồn nền	Được, nhưng lấy đi cả phần cao của giọng, để lại tiếng "lấp lánh"	Khử sâu hơn nhiều mà giọng gần như không suy suyển
Khử tiếng vọng phòng	Gần như không làm được	Làm được, và đây là bước nhảy lớn nhất của cả nhóm
Khử tiếng ù, tiếng rít máy	Được bằng bộ lọc khe hẹp, nhưng phải dò tay từng tần số	Tự dò và khử, kể cả khi tần số trôi
Tách lời khỏi nhạc nền	Không làm được	Làm được ở mức dùng trong dựng; chưa đủ để phát hành riêng
Khớp không gian giữa hai lần thu	Làm tay bằng hồi âm nhân tạo, rất khó khớp	Làm được tự động, tiết kiệm nhiều ngày công lồng tiếng

12.12.2 Vì sao nhóm này gần như không còn tranh cãi

Vì nó không tạo ra thứ gì mới. Nó chỉ lấy đi thứ không mong muốn ra khỏi một bản thu có thật. Không có nội dung nào bị bịa ra, không có người nào bị mô phỏng, nên không phát sinh nghĩa vụ nào ở Phần D. Nối lại mục 10.23 của Chương 10: các bước dọn tiếng nêu ở đó vẫn đúng nguyên về thứ tự; chỉ có công cụ ở từng bước là đổi.

Hình 12.23 so sánh phổ tần của cùng một đoạn thoại trong cửa sổ Frequency Analysis trước và sau khi giảm tạp âm theo dấu vân tay; sàn tạp âm ở vùng cao hạ xuống rõ rệt, còn dải thoại quanh vài trăm hertz gần như giữ nguyên.



Hình 12.23 Phổ tần của tệp thoại có tạp âm hồng mẫu tổng hợp, trước (trái) và sau (phải) khi giảm tạp âm trong Adobe Audition

Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Audition 2026, bản 26.5), 07/11/2026.

CẢNH BÁO 12.2 — MỘT GIỚI HẠN VẪN PHẢI NHỚ

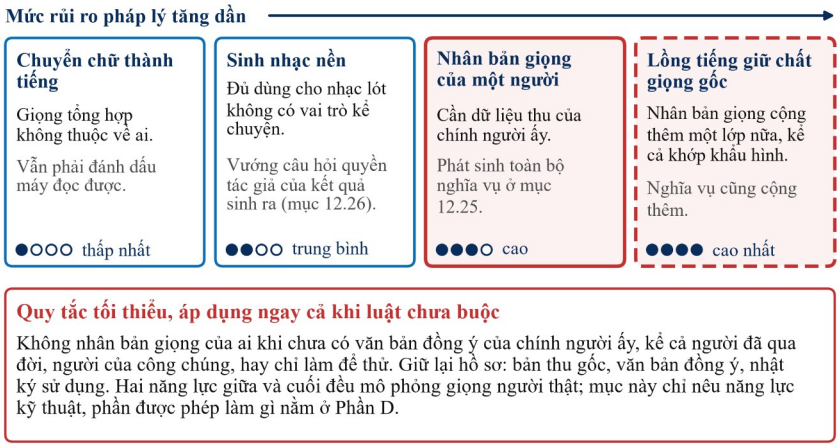
Tách lời khỏi nhạc nền cho ra sản phẩm nghe được, nhưng nghe kỹ vẫn còn dấu vết: phần cao của nhạc bám vào giọng, và phần trầm của giọng bị hụt.

Dùng bản tách để dựng thì tốt. Dùng bản tách để phát hành thì phải nghe lại trên hệ thống lớn trước khi quyết định.

12.13 Hậu kỳ tiếng — giọng nhân bản và lồng tiếng

12.13.1 Bốn năng lực kỹ thuật

Mục này chỉ nêu năng lực kỹ thuật. Toàn bộ phần quyền nhân thân, yếu tố thật sự quyết định có được dùng hay không, nằm ở mục 12.25, và người đọc nên đọc hai mục này liền nhau.



Hình 12.24 Bốn năng lực của nhóm giọng nhân bản, xếp theo mức rủi ro pháp lý

- **Chuyển chữ thành tiếng nói.** Giọng tổng hợp không thuộc về ai nên rủi ro pháp lý thấp nhất trong nhóm, nhưng vẫn thuộc diện phải đánh dấu máy đọc được.
- **Nhân bản giọng của một người.** Việc này cần dữ liệu thu của chính người ấy, và toàn bộ nghĩa vụ ở mục 12.25 phát sinh từ đây.
- **Lồng tiếng sang ngôn ngữ khác, giữ chất giọng gốc.** Kỹ thuật đã đủ dùng, kể cả khớp khẩu hình, nhưng vì đây là nhân bản giọng cộng thêm một lớp xử lý nữa nên nghĩa vụ cũng tăng thêm.
- **Sinh nhạc nền.** Công cụ đủ dùng cho nhạc lót không có vai trò kể chuyện, song vẫn vướng câu hỏi ở mục 12.26 về quyền tác giả của kết quả sinh ra.

12.13.2 Vì sao đây là nhóm nhiều rủi ro nhất

Vì giọng nói là một phần của nhân thân, không phải một loại tư liệu. Một cảnh nền bị sinh sai thì sửa lại; một giọng bị nhân bản mà không có sự đồng ý thì đã xâm phạm quyền của một con người cụ thể, và việc xoá tệp đi không xoá được việc đã xảy ra.

> **QUY TẮC 12.6 — QUY TẮC TỐI THIỂU, ÁP DỤNG NGAY CẢ KHI LUẬT CHƯA BUỘC**
>
> Không nhân bản giọng của bất kỳ ai khi chưa có văn bản đồng ý của chính người ấy, kể cả người đã qua đời, nhân vật của công chúng, hoặc khi chỉ làm để thử.
>
> Văn bản đồng ý phải đủ sáu điểm ở mục 12.25. Một câu "tôi đồng ý" trong tin nhắn không phải là văn bản đồng ý.
>
> Cần giữ lại toàn bộ hồ sơ gồm bản thu gốc, văn bản đồng ý và nhật ký sử dụng, vì đây là căn cứ duy nhất bảo vệ người làm nghề khi có tranh chấp.

12.14 Quyết định dùng hay không — và bài thực hành số 2

12.14.1 Chi phí thật cho một kết quả dùng được

Mục 12.5 đã lập công thức thời gian; quy ra tiền thì được công thức song song:

$$C_{\text{eff}} = \frac{C_{\text{gen}} + C_{\text{hum}}}{\eta} \tag{12.5}$$

trong đó C_{gen} là chi phí máy sinh một lần và C_{hum} là chi phí công người xem và quyết định một lần. Công người được quy ra tiền theo thời gian xem và đơn giá giờ.

12.14.2 Điểm hoà vốn

Đầu tư vào một công cụ mới bao giờ cũng có một khoản trả trước: tiền giấy phép, thời gian học nghề, công dựng lại quy trình. Khoản ấy chỉ đáng bỏ ra nếu việc sẽ phải làm đủ nhiều lần.

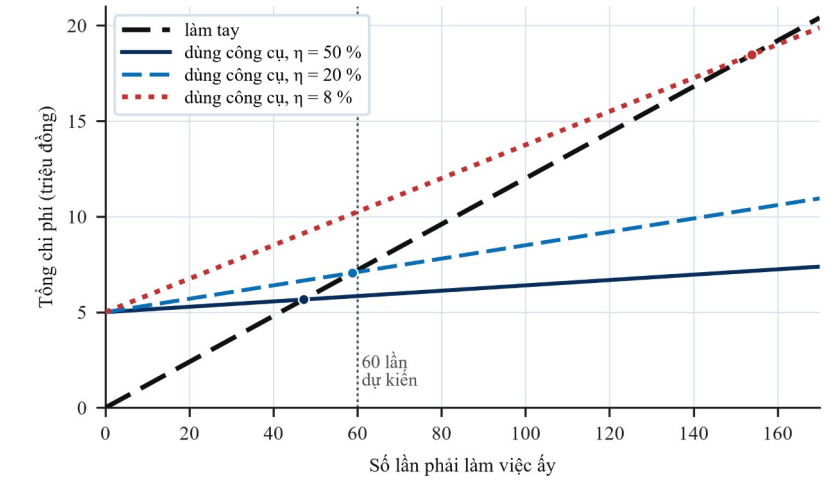
$$N^{*} = \frac{C_{\text{setup}}}{C_{\text{man}} - C_{\text{eff}}} \quad , \quad C_{\text{man}} > C_{\text{eff}} \tag{12.6}$$

Trong đó C_{setup} là chi phí thiết lập một lần, C_{man} là chi phí làm tay một lần, và N* là số lần phải làm để hoà vốn. Điều kiện $C_{man} > C_{eff}$ là điều kiện tồn tại: nếu công cụ mới không rẻ hơn làm tay thì không có điểm hoà vốn nào, và "không đáng dùng" là một kết luận hợp lệ.

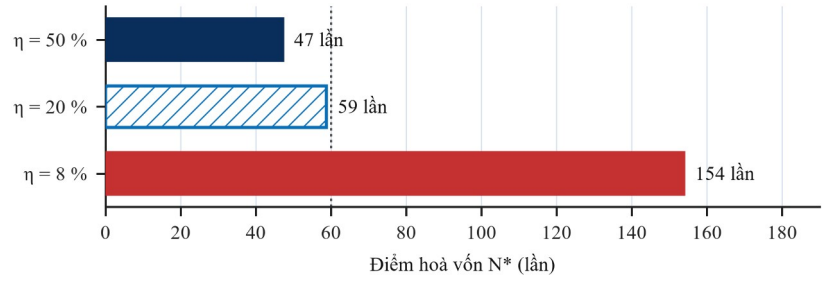
Bảng 12.7 Cùng một công cụ, ba tỉ lệ dùng được, ba kết luận khác nhau

Tỉ lệ dùng được η	Chi phí thật một kết quả C_{eff} (đồng)	Điểm hoà vốn N* (lần)	Kết luận với 60 lần dự kiến
50 %	14.000	47	đáng đầu tư
20 %	35.000	59	đáng đầu tư
8 %	87.500	154	chưa đáng

Bảng trên dùng chung một bộ giả thiết: chi phí thiết lập 5.000.000 đồng, làm tay 120.000 đồng một lần, máy sinh 2.000 đồng một lần, người xem 120 giây với đơn giá 150.000 đồng một giờ, tức 5.000 đồng một lần xem; chỉ có η thay đổi.



(a) Tổng chi phí; chấm tròn là điểm hoà vốn N*

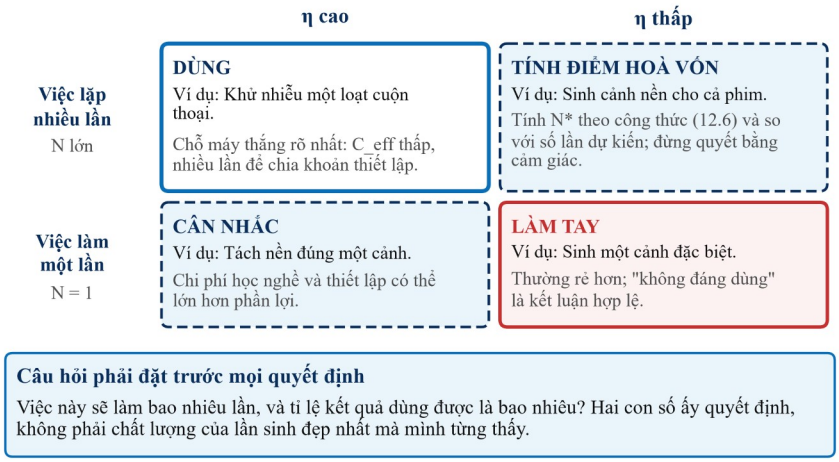


(b) Số lần phải làm mới hoà vốn, so với 60 lần dự kiến

Ví dụ minh hoạ, không phải số đo thật: chi phí thiết lập 5.000.000 đồng; làm tay 120.000 đồng một lần; máy sinh 2.000 đồng một lần; người xem 120 giây với đơn giá 150.000 đồng một giờ, tức 5.000 đồng một lần xem. Chỉ η thay đổi.

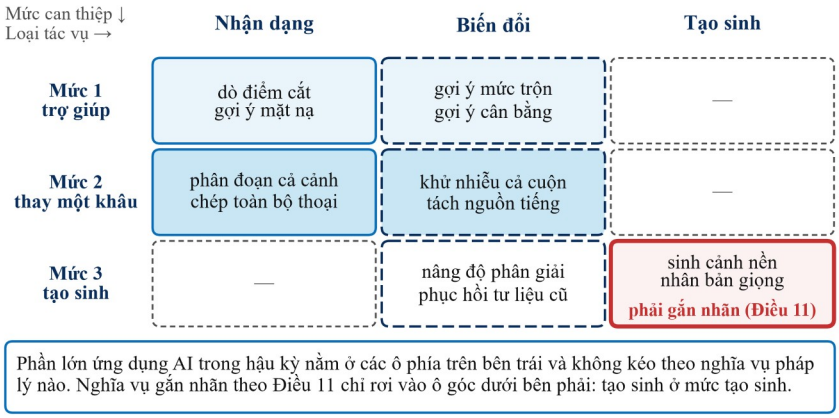
Hình 12.25 Hai đường chi phí và điểm hoà vốn, vẽ cho ba giá trị của tỉ lệ dùng được

Điều mà bảng và hình cùng nói: η không phải một chi tiết kỹ thuật. Nó là con số duy nhất làm kết luận đổi chiều. Cùng một công cụ, cùng một giá, mà η thay đổi thì lời khuyên đầu tư đảo ngược hoàn toàn. Và η thì không tra được ở đâu — phải tự đo.



Hình 12.26 Bảng quyết định: bốn tổ hợp của tỉ lệ dùng được và số lần phải làm

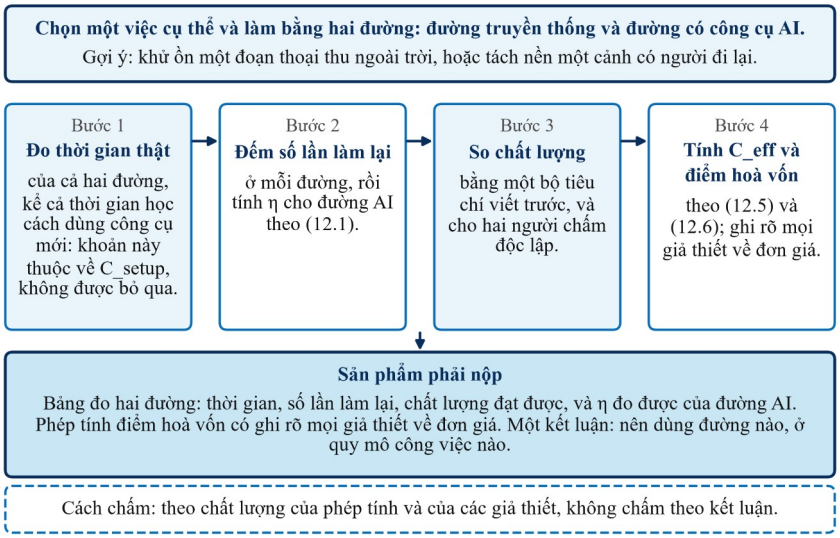
12.14.3 Đọc lại cả Phần B bằng hai trục



Hình 12.27 Toàn bộ Phần B đặt lên hai trục: ba loại tác vụ và ba mức can thiệp

Bảng trên gộp hai trục đã dựng ở Phần A. Đọc nó thì thấy ngay vì sao phần lớn ứng dụng AI trong hậu kỳ không kéo theo nghĩa vụ pháp lý nào: chúng nằm ở các ô phía trên bên trái. Nghĩa vụ gắn nhãn theo Điều 11 chỉ rơi vào ô góc dưới bên phải — tạo sinh ở mức tạo sinh — và chỉ ô ấy.

12.14.4 Bài thực hành số 2 — làm một khâu bằng hai đường



Hình 12.28 Bài thực hành số 2 — bốn bước và kết quả cần có

Người đọc chọn một việc cụ thể và thực hiện bằng hai đường: đường truyền thống và đường có công cụ AI. Việc có thể chọn là khử ồn một đoạn thoại thu ngoài trời, hoặc tách nền một cảnh có người đi lại.

- **Đo thời gian thực của cả hai đường,** kể cả thời gian học cách dùng công cụ mới, vì khoản này thuộc về C_{setup} và không được bỏ qua.
- **Đếm số lần phải làm lại ở mỗi đường** và tính η cho đường AI.
- **So sánh chất lượng bằng một bộ tiêu chí viết trước,** với hai người đánh giá độc lập.
- **Tính C_{eff} và điểm hoà vốn,** ghi rõ mọi giả thiết về đơn giá.

> **THỰC HÀNH 12.2 — CÁCH TỰ ĐÁNH GIÁ BÀI 2**
>
> Đánh giá theo chất lượng của phép tính và của các giả thiết, không theo kết luận.
>
> Khi hai người cùng làm bài đưa ra hai kết luận trái ngược nhưng cùng lập luận chặt chẽ thì cả hai đều có giá trị, và nên dành thời gian tìm ra chỗ khác nhau giữa các giả thiết của mỗi bên.

12.15 Sản xuất ảo — phần dựng cảnh thời gian thực

12.15.1 Nối tiếp Chương 11 ở chỗ nào

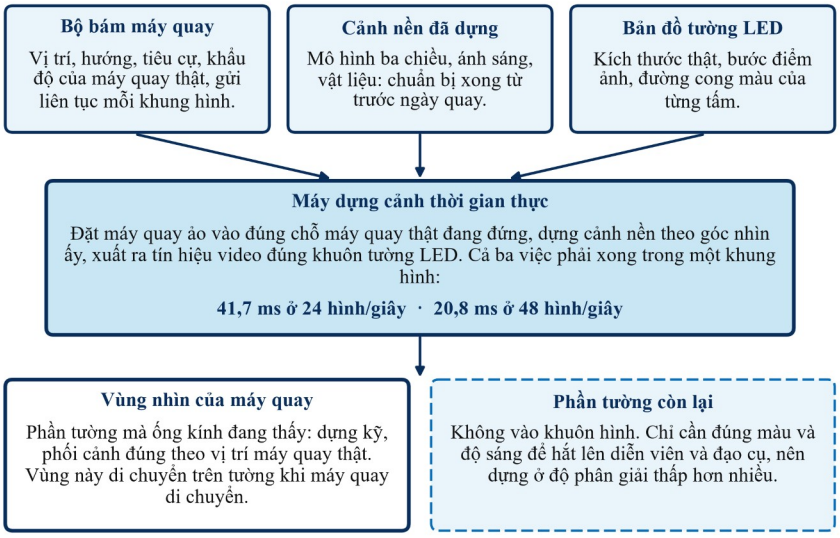
Các mục 11.24 tới 11.26 của Chương 11 đã trình bày tường LED như một thiết bị hiển thị, với bước điểm ảnh, độ chói, tần số quét và đồng bộ với cửa chập máy quay, tức là phần cứng. Mục này trình bày phần còn lại: nội dung chạy trên tường ấy và cách nội dung đó được dựng.

Hình 12.29 cho thấy một trường quay sản xuất ảo cỡ nhỏ của cơ sở đào tạo, với tường LED cong hiển thị cảnh nền dựng bằng máy tính và giàn treo đèn phía trên.



Hình 12.29 Trường quay sản xuất ảo cỡ nhỏ: tường LED cong hiển thị cảnh nền dựng thời gian thực, giàn treo đèn phía trên (Đại học Portsmouth)

Nguồn: Tim Sheerman-Chase, CC BY 2.0, Wikimedia Commons (truy cập 05/11/2026).



Hình 12.30 Một khung hình sản xuất ảo được dựng như thế nào — ba nguồn đầu vào và hai vùng trên tường

12.15.2 Ba nguồn đầu vào

- **Bộ bám máy quay.** Thiết bị đo vị trí, hướng, tiêu cự và khẩu độ của máy quay thật rồi gửi liên tục theo từng khung hình; thiết bị này nối thế giới thật với thế giới dựng nên không thể thiếu trong sản xuất ảo.
- **Cảnh nền đã dựng.** Mô hình ba chiều, ánh sáng, vật liệu — chuẩn bị xong từ trước ngày quay. Đây là chỗ chứa gần như toàn bộ khối lượng công việc, và là chủ đề của mục 12.16.
- **Bản đồ tường LED.** Gồm kích thước thật, bước điểm ảnh và đường cong màu của từng tấm, lấy từ phép đo ở Chương 11; máy dựng cần biết vị trí tường thật trong không gian thì mới đặt được máy quay ảo cho đúng.

12.15.3 Vùng nhìn của máy quay và phần tường còn lại

Đây là mẹo tiết kiệm cốt lõi của cả kỹ thuật. Máy dựng chỉ dựng kỹ phần tường mà ống kính đang thấy — gọi là vùng nhìn của máy quay — với phối cảnh đúng theo vị trí máy quay thật. Phần tường còn lại chỉ cần đúng màu và độ sáng để hắt lên diễn viên và đạo cụ, nên dựng ở độ phân giải thấp hơn nhiều [7].

Hệ quả cần nhớ: khi máy quay di chuyển, vùng nhìn ấy di chuyển trên tường theo. Nếu bộ bám máy quay trễ, vùng nhìn chạy sau máy quay và khán giả thấy nền trượt so với tiền cảnh. Đây là lỗi thường gặp nhất ở trường quay ảo, và là lý do vai trò kỹ thuật viên bám máy quay ở mục 12.16.4 tồn tại.

> **GHI NHỚ 12.3 — BÀI TOÁN THỜI GIAN, NỐI LẠI VỚI CHƯƠNG 11**
>
> Cả ba việc — đặt máy quay ảo, dựng cảnh nền theo góc nhìn ấy, xuất ra tín hiệu đúng khuôn tường — phải xong trong một khung hình.
>
> Ở 24 hình một giây, thời lượng đó là khoảng 41,7 mili giây, còn ở 48 hình một giây chỉ còn khoảng 20,8 mili giây.
>
> Đây là lý do sản xuất ảo cần máy tính mạnh không phải để đẹp hơn, mà để kịp. Cảnh nền dựng đẹp tới đâu mà không kịp thì vô dụng.

12.16 ICVFX — quy trình và ê-kíp mới

12.16.1 Kỹ xảo hoàn thành ngay trong máy quay

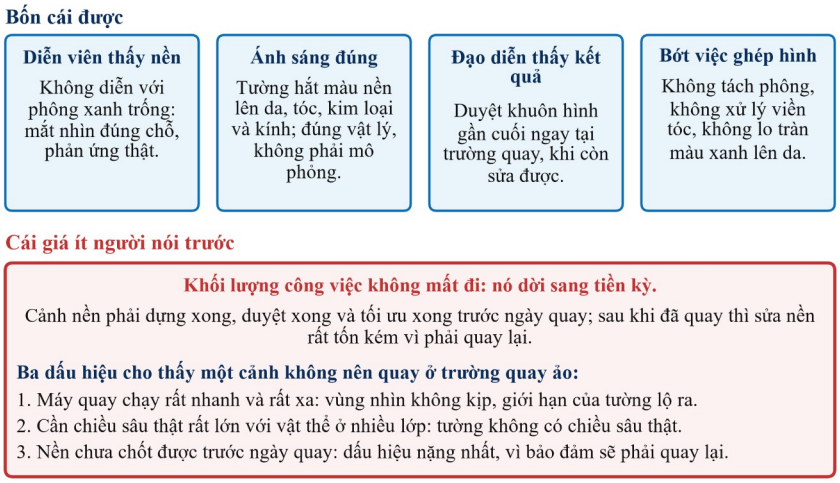
Tên gọi của phương pháp này nói đúng điều cốt lõi: kỹ xảo được hoàn thành ngay trong máy quay, chứ không ghép vào sau. Cái mà máy quay ghi được đã là khuôn hình gần cuối cùng.

Hình 12.31 cho thấy một bộ sản xuất ảo trưng bày, trong đó cảnh sân bóng dựng thời gian thực hiển thị trên tường LED, cánh tay robot mang máy quay và các màn hình giám sát phía trước.



Hình 12.31 Bộ sản xuất ảo trưng bày tại triển lãm BIRTV 2025: cảnh nền trên tường LED, cánh tay robot mang máy quay và bàn giám sát

Nguồn: TurnOnTheNight, CC BY-SA 4.0, Wikimedia Commons (truy cập 05/11/2026).



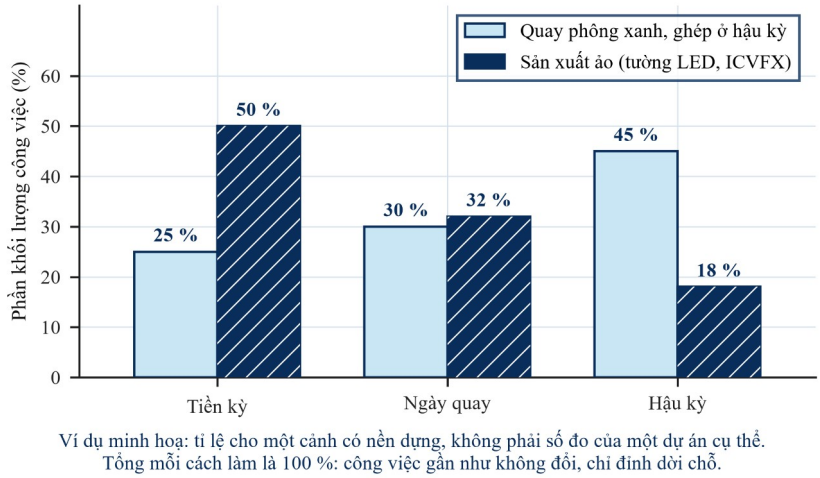
Hình 12.32 Bốn cái được của ICVFX và cái giá ít người nói trước

12.16.2 Bốn cái được

- Diễn viên thấy nền: không phải diễn với một tấm phông xanh trống. Mắt nhìn đúng chỗ, phản ứng thật. Với cảnh có cảm xúc, khác biệt này lớn hơn mọi khác biệt kỹ thuật.
- Ánh sáng đúng: tường hắt màu nền lên da, lên tóc, lên bề mặt kim loại và kính. Đây là thứ tốn công nhất khi ghép hình ở hậu kỳ, và ở đây nó có sẵn, đúng vật lý, không phải mô phỏng.
- Đạo diễn thấy kết quả: duyệt khuôn hình gần cuối ngay tại trường quay. Quyết định được đưa ra khi còn sửa được, thay vì ba tháng sau khi mọi người đã giải tán.
- Giảm việc ghép hình: không phải tách phông, không phải xử lý viền tóc và không lo màu xanh tràn lên da diễn viên.

12.16.3 Cái giá ít người nói trước

Khối lượng công việc không mất đi. Nó dời sang tiền kỳ. Cảnh nền phải dựng xong, duyệt xong và tối ưu xong trước ngày quay; sau khi đã quay thì sửa nền rất tốn kém vì phải quay lại.



Hình 12.33 Khối lượng công việc dời từ hậu kỳ sang tiền kỳ

Hình trên là hình quan trọng nhất của mục này, và con số trong đó là tỉ lệ minh hoạ chứ không phải số đo của một dự án cụ thể. Điều cần nhớ là hình dáng: tổng công việc gần như không đổi, nhưng đỉnh dời từ hậu kỳ về tiền kỳ. Hệ quả quản lý rất nặng — tiền và người phải được rót sớm hơn nhiều so với thói quen của một đoàn phim, mà thói quen ấy thì đã hình thành qua hàng chục năm.

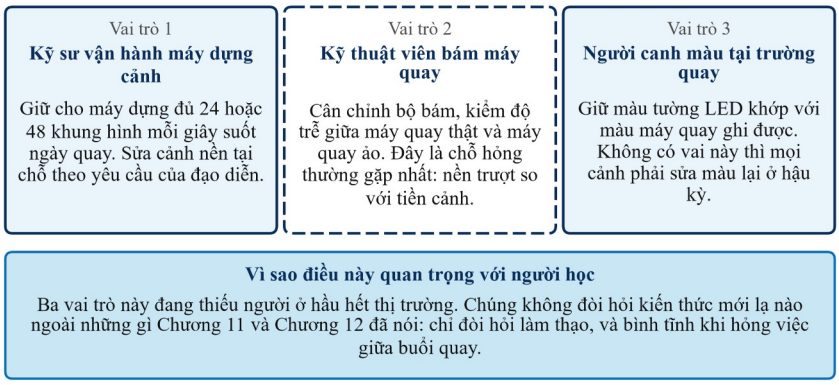
CẢNH BÁO 12.3 — BA DẤU HIỆU CHO THẤY MỘT CẢNH KHÔNG NÊN QUAY Ở TRƯỜNG QUAY ẢO

Cảnh có máy quay di chuyển rất nhanh và rất xa, khi vùng nhìn không kịp theo và giới hạn của tường lộ ra.

Cảnh cần chiều sâu thật rất lớn với vật thể ở nhiều lớp, vì tường là một mặt phẳng hoặc mặt cong và không có chiều sâu thật.

Cảnh có nền chưa chốt được trước ngày quay; đây là dấu hiệu nghiêm trọng nhất vì gần như chắc chắn sẽ phải quay lại.

12.16.4 Ba vai trò mới trong ê-kíp



Hình 12.34 Ba vai trò mà ê-kíp truyền thống không có

Ba vai trò ở hình trên đang thiếu người ở hầu hết thị trường. Điều đáng lưu ý với người đọc là các vai trò này không đòi hỏi kiến thức nào ngoài những gì Chương 11 và Chương 12 đã trình bày, mà đòi hỏi sự thành thạo và khả năng bình tĩnh khi sự cố xảy ra giữa buổi quay, lúc cả trăm người đang chờ.

Hình 12.35 cho thấy bàn điều khiển của một trường quay sản xuất ảo, nơi những người vận hành theo dõi cảnh nền và máy quay trong khi việc quay diễn ra ở vùng sân khấu phía sau.



Hình 12.35 Bàn điều khiển và vùng quay của một trường quay sản xuất ảo: người vận hành ngồi tại bàn, máy quay và tường LED ở phía sau

Nguồn: Tim Sheerman-Chase, CC BY 2.0, Wikimedia Commons (truy cập 05/11/2026).

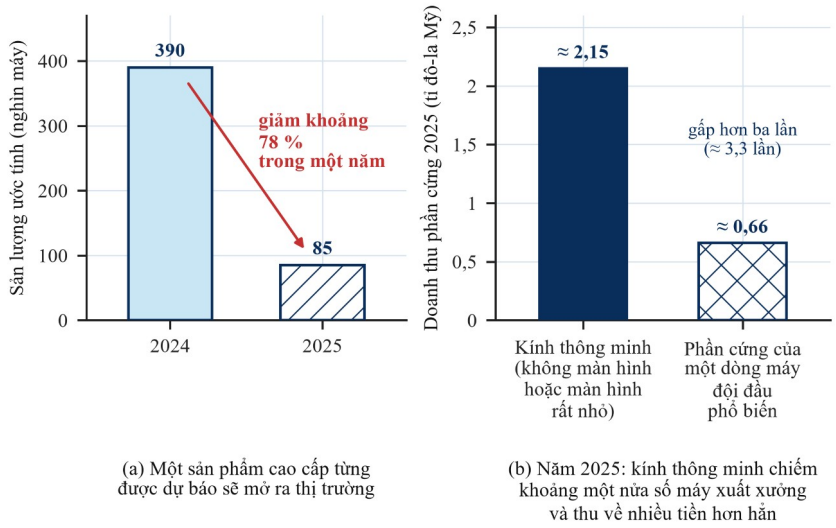
12.17 XR và các phương tiện mới — đọc đúng hiện trạng

12.17.1 Vì sao mục này không có bảng thiết bị

Các bảng thiết bị thực tế mở rộng nhanh chóng lỗi thời: một bảng bảy thiết bị, sau chừng mười tám tháng, có thể có một dòng liệt kê sản phẩm mà nhà sản xuất đã ngừng làm và không có sản phẩm kế nhiệm. Đây không phải lỗi của người viết bảng — đây là điều tất yếu xảy ra với mọi bảng thiết bị trong mọi cuốn sách.

Vì vậy mục này không đưa bảng thiết bị mà trình bày một nội dung khác: cách đọc một phương tiện mới. Danh mục thiết bị hiện hành được đưa vào phụ lục cuối sách, tách hẳn khỏi phần nguyên lý, để có thể thay hằng năm mà không phải sửa nội dung nào của chương.

12.17.2 Một ví dụ đáng để học thuộc



(a) Một sản phẩm cao cấp từng được dự báo sẽ mở ra thị trường

(b) Năm 2025: kính thông minh chiếm khoảng một nửa số máy xuất xưởng và thu về nhiều tiền hơn hẳn

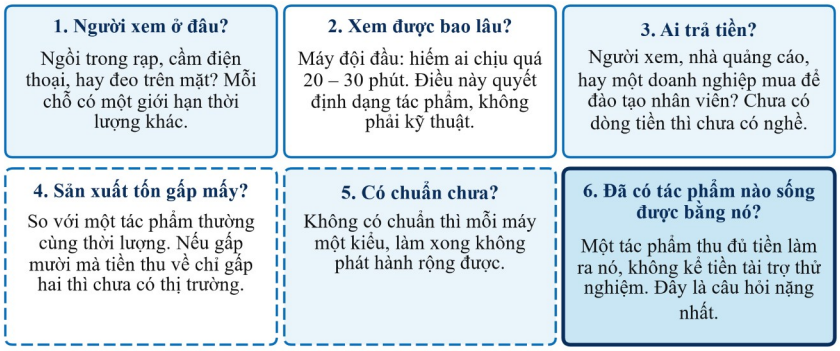
Hình 12.36 Thị trường thiết bị đội đầu rẽ hướng: sản lượng một sản phẩm cao cấp và cơ cấu doanh thu năm 2025

Câu chuyện ở hình trên đáng để đọc kỹ. Khoảng năm 2024, dự báo phổ biến trong ngành là kính thực tế ảo cao cấp sẽ mở ra một loại nội dung mới và một thị trường mới. Dự báo ấy dựa trên những nhận định đúng về năng lực kỹ thuật: máy đội đầu quả thật đã nhẹ hơn, hiển thị quả thật đã tốt hơn nhiều, và quả thật đã có những chỗ dùng thật.

Nhưng tới 2025 – 2026, thị trường rẽ sang một hình thức sản phẩm khác hẳn: kính thông minh — nhẹ, không màn hình hoặc màn hình rất nhỏ, đeo được cả ngày, giá bằng một phần mười. Loại này chiếm khoảng một nửa số máy xuất xưởng trong nhóm thiết bị đội đầu, và thu về nhiều tiền hơn hẳn. Còn sản phẩm cao cấp từng được kỳ vọng thì sản lượng rơi khoảng 78 phần trăm trong một năm và việc sản xuất bị dừng.

> **GHI NHỚ 12.4 — BÀI HỌC, PHÁT BIỂU Ở DẠNG DÙNG LẠI ĐƯỢC**
>
> **Dự báo về năng lực thường đúng:** kỹ thuật đi theo đường đã thấy.
>
> Dự báo về hình thức sản phẩm sẽ thắng thì thường sai, vì hình thức do người dùng chọn, mà người dùng không chọn theo thông số kỹ thuật.
>
> Do đó, người làm nghề không nên đặt cược sự nghiệp vào một hình thức sản phẩm mà nên đầu tư vào năng lực, vốn chuyển được sang hình thức khác khi thị trường thay đổi.

12.17.3 Sáu câu hỏi để đọc một phương tiện mới



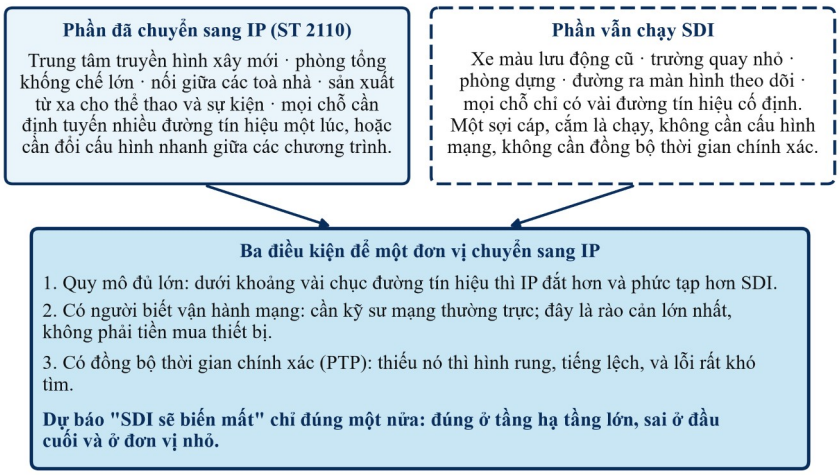
Hình 12.37 Sáu câu hỏi để đọc một phương tiện mới

Sáu câu hỏi này áp dụng được cho một phương tiện chưa ra đời. Trong đó, câu cuối cùng (đã có tác phẩm nào sống được bằng phương tiện ấy chưa) là câu có trọng lượng lớn nhất nhưng ít khi được trả lời thẳng. Một tác phẩm thu đủ tiền làm ra nó, không tính tiền tài trợ thử nghiệm, là bằng chứng duy nhất cho thấy một phương tiện đã trở thành một nghề thực sự chứ chưa phải một cuộc thí nghiệm kéo dài.

12.18 Hạ tầng — sản xuất trên nền IP và trên đám mây

12.18.1 Hiện trạng: lai ghép, không phải thay hẳn

Truyền tín hiệu video chuyên nghiệp trên hạ tầng mạng theo chuẩn ST 2110 đã trở thành xương sống của các cơ sở lớn. Tuy nhiên, dự báo phổ biến vài năm trước, rằng SDI sẽ biến mất, chỉ đúng một nửa [5].



Hình 12.38 Hạ tầng lai ghép: phần đã chuyển sang IP, phần vẫn chạy SDI, và ba điều kiện để chuyển

Ba điều kiện ở hình trên đáng nhắc lại, vì điều kiện thứ hai thường bị bỏ qua khi lập dự toán: rào cản lớn nhất không phải tiền mua thiết bị mà là có người biết vận hành mạng thường trực. Một hệ thống ST 2110 không có kỹ sư mạng là một hệ thống sẽ hỏng vào đúng lúc không được phép hỏng.

Với người đọc, kết luận thực tế là sẽ gặp cả hai công nghệ, thường trong cùng một toà nhà. Kỹ năng cần có không phải là chọn một bên mà là biết vị trí nối giữa hai hệ thống và các dạng hỏng thường gặp ở đó.

12.18.2 Dựng và duyệt trên đám mây — bài toán độ trễ

Làm việc từ xa trên tư liệu đặt ở trung tâm dữ liệu đặt ra một bài toán mà mười một chương trước chưa đề cập: độ trễ vòng từ lúc bấm phím tới lúc thấy kết quả trên màn hình.

$$t_{rt} = t_{net} + t_{dec} + t_{disp} \quad [\text{ms}] \tag{12.7}$$

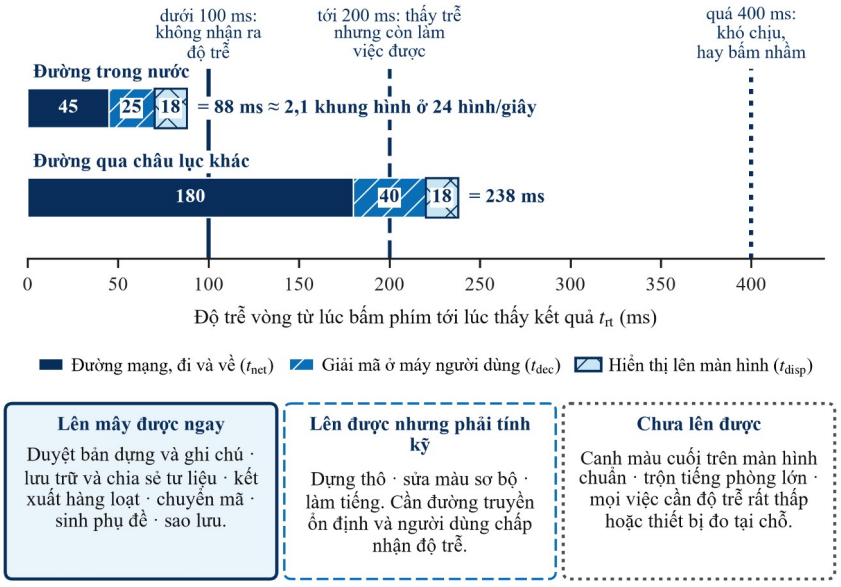
Trong đó t_{net} là thời gian đi và về trên đường mạng, t_{dec} là thời gian giải mã ở máy người dùng, và t_{disp} là thời gian màn hình hiển thị. Ba chặng này cộng dồn vào nhau và chặng nào cũng đáng kể.

VÍ DỤ 12.2 — VÍ DỤ TÍNH

Với đường trong nước, mạng 45 ms, giải mã 25 ms, hiển thị 18 ms thì t_{rt} = 88 ms, tức khoảng 2,1 khung hình ở 24 hình một giây; giá trị này nằm dưới ngưỡng 100 ms nên người dựng không nhận ra độ trễ.

Với đường qua châu lục khác, mạng 180 ms, giải mã 40 ms thì t_{rt} = 238 ms, đã vượt ngưỡng 200 ms; người dựng vẫn xem duyệt được nhưng không còn bấm dừng đúng khung hình.

Các ngưỡng cần ghi nhớ: dưới 100 ms thì không nhận ra độ trễ; tới 200 ms thì thấy trễ nhưng vẫn làm việc được; quá 400 ms thì gây khó chịu và dễ bấm nhầm.



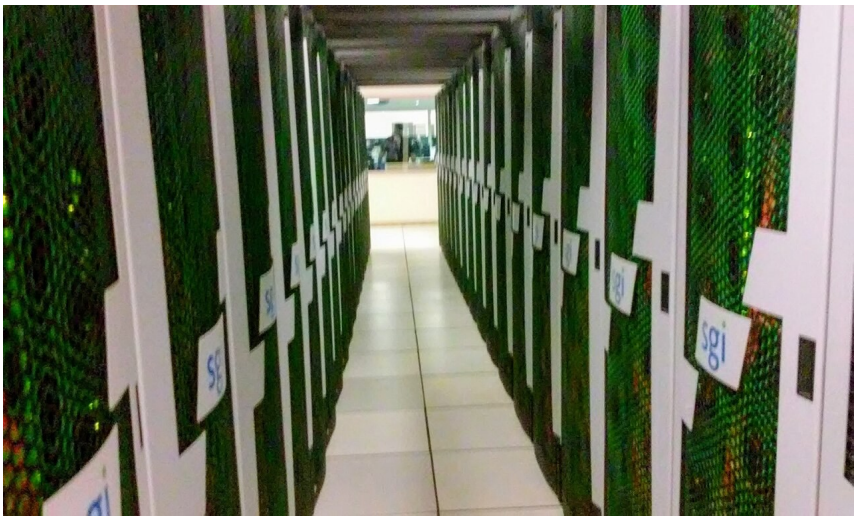
Hình 12.39 Ngân sách độ trễ của một buổi dựng từ xa, và ba nhóm công việc theo mức đưa được lên đám mây

Ba ngăn ở phần phải của hình không cố định theo thời gian. Ngăn giữa đang dần chuyển lên ngăn trên khi đường truyền tốt lên. Còn ngăn dưới — canh màu cuối trên màn hình chuẩn, trộn tiếng phòng lớn — thì mười năm nay chưa nhúc nhích, và lý do không phải kỹ thuật mạng mà là: những việc ấy cần một căn phòng đã được hiệu chuẩn, mà căn phòng thì không đưa lên mạng được.

12.18.3 Băng thông và dung lượng — nối lại các phép tính đã có

Hai bài toán còn lại của hạ tầng đã được lập công thức ở chỗ khác của cuốn sách nên ở đây chỉ nêu nơi tra cứu: lưu lượng dữ liệu của một luồng tín hiệu tính theo mục 10.5, và dung lượng kho khi sinh nhiều phương án tính theo công thức ở mục 12.11.3.

Hình 12.40 cho thấy một dãy tủ máy chủ trong trung tâm tính toán, loại hạ tầng nơi các công việc dựng, nén và suy luận trên đám mây thực sự chạy.



Hình 12.40 Dãy tủ máy chủ của một siêu máy tính (Pleiades) trong trung tâm tính toán

Nguồn: Oleg Alexandrov, CC BY-SA 4.0, Wikimedia Commons (truy cập 05/11/2026).

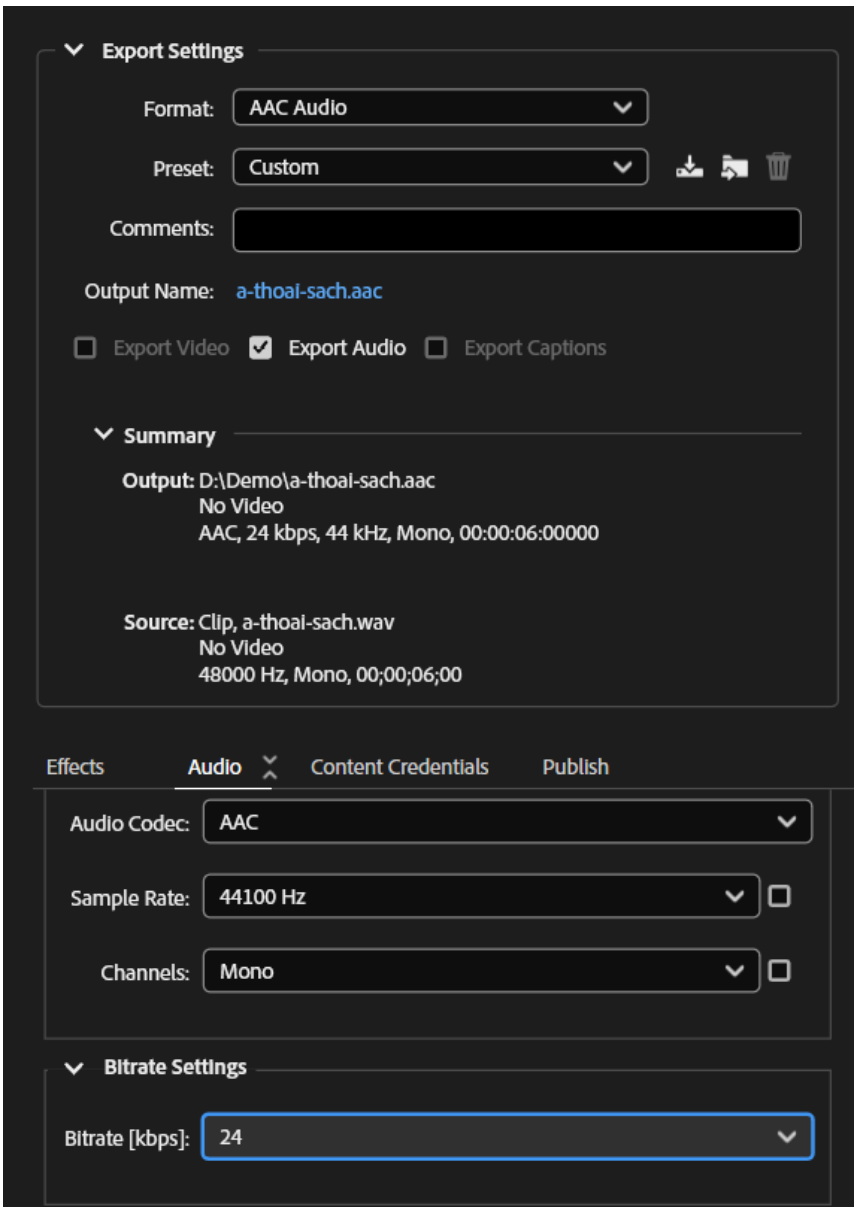
Một điểm nối cần lưu ý khi hai bài toán ấy gặp nhau trên đám mây: chi phí đưa dữ liệu ra khỏi trung tâm dữ liệu thường cao hơn chi phí lưu, và cao hơn nhiều so với ước lượng ban đầu. Một kho cao điểm hàng trăm gi-ga-bai không đắt khi nằm yên, nhưng đắt ngay khi phải kéo về máy tại chỗ nhiều lần. Nên khi lập dự toán, phải đếm số lần dữ liệu đi ra, không chỉ đếm dung lượng.

12.19 Hướng nén mới dùng mạng thần kinh

12.19.1 Nguyên lý: bộ mã hoá học được thay bộ mã hoá thiết kế tay

Mọi bộ mã hoá mà Chương 4 và Chương 8 đã đề cập đều được con người thiết kế từng bước gồm chia khối, dự đoán, biến đổi, lượng tử hoá và mã hoá entropy. Mỗi bước đều giải thích được, đo được và viết được thành chuẩn trên giấy.

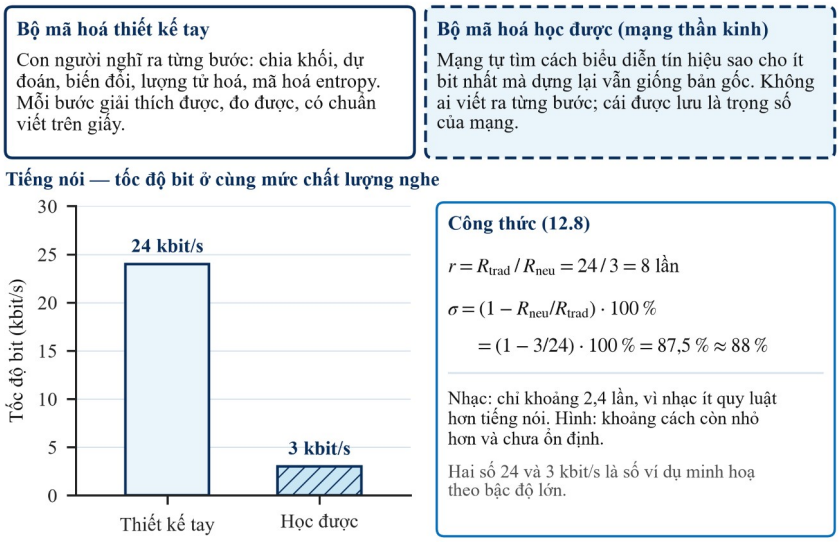
Hình 12.41 cho thấy một bộ mã hoá cổ điển được đặt ở tốc độ bit rất thấp: Adobe Media Encoder nén đoạn thoại đơn kênh bằng AAC ở 24 kbit/s, làm mốc để đối chiếu với các bộ mã hoá bằng mạng thần kinh.



Hình 12.41 Adobe Media Encoder đặt AAC (bộ mã hoá cổ điển, thiết kế thủ công) ở 24 kbit/s, đơn kênh, cho một đoạn thoại mẫu tổng hợp; đây là mốc đối chiếu với nén bằng mạng thần kinh

Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Media Encoder 2026), 07/11/2026.

Hướng mới thay toàn bộ chuỗi ấy bằng một mạng thần kinh tự tìm cách biểu diễn tín hiệu sao cho ít bit nhất mà dựng lại vẫn giống bản gốc. Không ai viết ra từng bước; cái được lưu là trọng số của mạng.

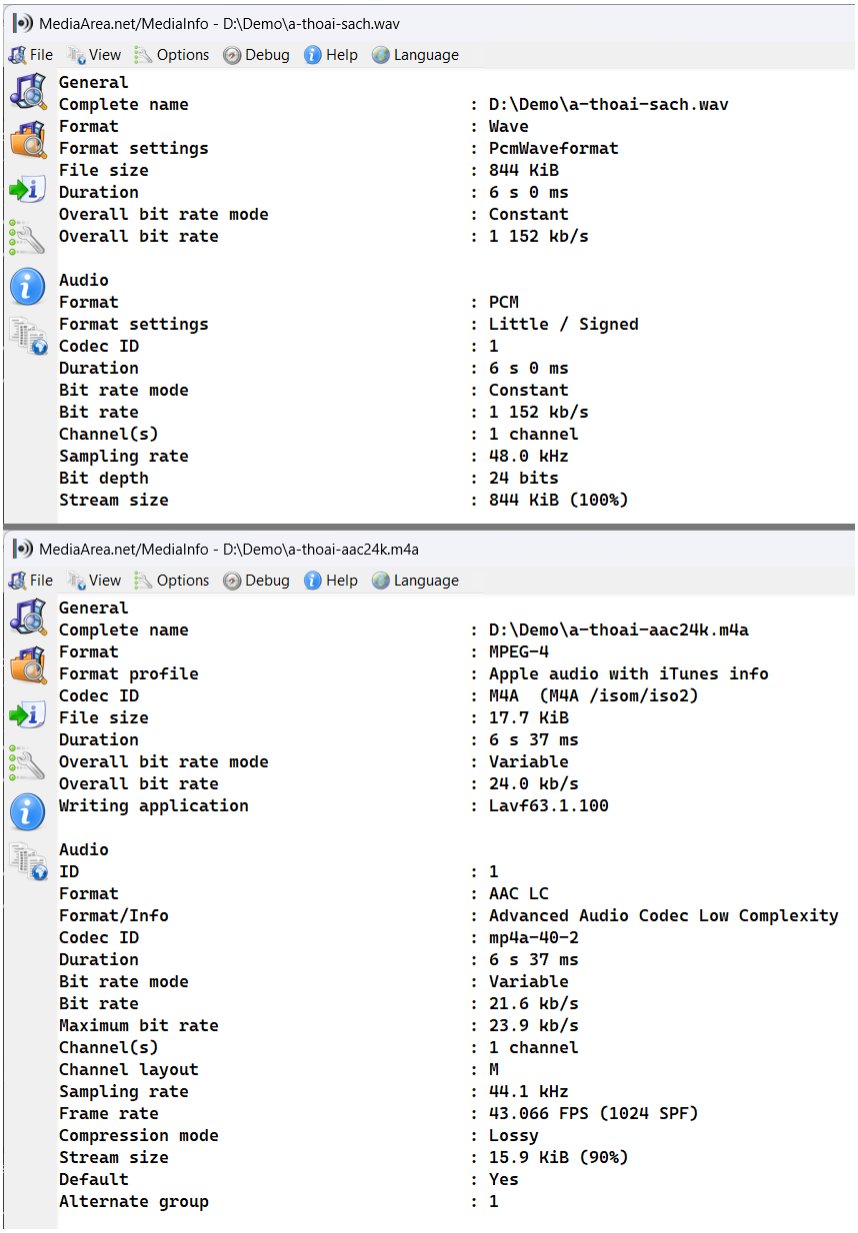


Hình 12.42 Hai đường mã hoá và mức nén đạt được ở tiếng nói

12.19.2 Vì sao nó thắng đậm ở tiếng nói

Tiếng nói là một tín hiệu rất có quy luật: nó được sinh ra bởi một bộ máy phát âm có cấu tạo giống nhau ở mọi người, và tuân theo một số hữu hạn cách. Một mạng học được quy luật ấy thì chỉ cần truyền phần sai khác so với cái nó đã biết, chứ không phải truyền toàn bộ dạng sóng.

Hình 12.43 đặt cạnh nhau thông tin MediaInfo của đoạn thoại gốc và của bản nén bằng AAC cổ điển: tốc độ bit giảm từ 1 152 kbit/s xuống 24 kbit/s, tức nhỏ hơn khoảng 48 lần, và là mốc mà các bộ mã hoá thần kinh phải vượt qua.



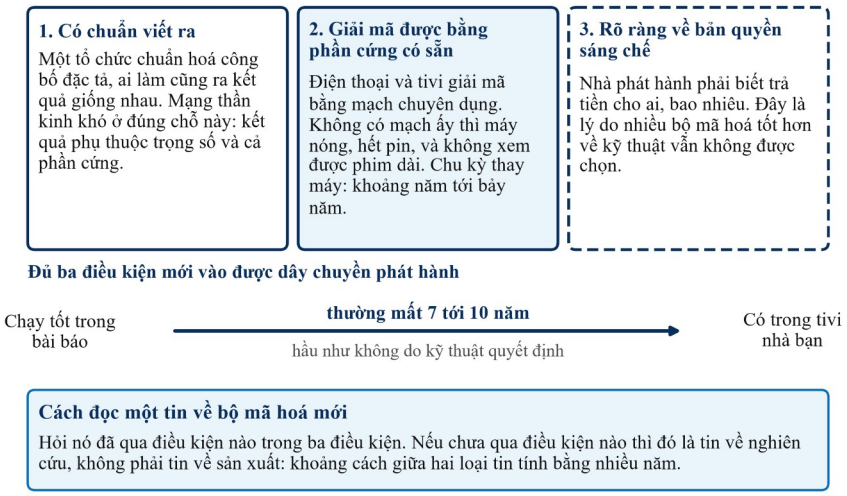
Hình 12.43 MediaInfo: đoạn thoại mẫu tổng hợp ở dạng PCM 1 152 kbit/s (trên) và sau khi nén bằng bộ mã hoá cổ điển AAC LC ở 24 kbit/s (dưới)

Nguồn: ảnh chụp màn hình do người soạn thực hiện (MediaInfo; tệp nén bằng Adobe Media Encoder 2026), 07/11/2026.

$$r = \frac{R_{\text{trad}}}{R_{\text{neu}}} \quad ; \quad \sigma = \left(1 - \frac{R_{\text{neu}}}{R_{\text{trad}}}\right) \cdot 100\ \% \tag{12.8}$$

Với hai con số minh hoạ ở hình trên: r = 8 lần, tức tiết kiệm σ = 88 % tốc độ bit ở cùng mức chất lượng nghe. Với nhạc thì lợi ích thấp hơn hẳn — khoảng 2,4 lần — vì nhạc ít quy luật hơn tiếng nói. Với hình thì khoảng cách còn nhỏ hơn nữa và chưa ổn định.

12.19.3 Vì sao tới nay nó vẫn chưa vào dây chuyền phát hành



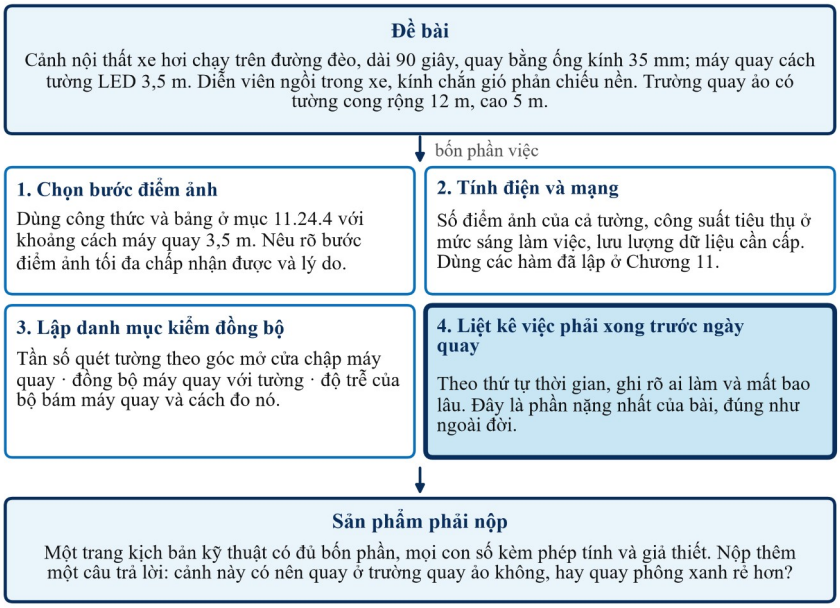
Hình 12.44 Ba điều kiện để một bộ mã hoá mới thật sự được dùng

Ba điều kiện ở hình trên giải thích một hiện tượng mà người đọc sẽ gặp lại nhiều lần trong nghề: khoảng cách giữa "chạy tốt trong bài báo" và "có trong tivi của người xem" thường là bảy tới mười năm, và khoảng cách ấy hầu như không do yếu tố kỹ thuật quyết định.

Nối lại mục 11.23.3 của Chương 11, nơi đã trình bày vì sao một chuẩn hiển thị mới mất nhiều năm mới phổ biến: lý do ở đây tương tự, và điều kiện thứ hai (giải mã được bằng phần cứng có sẵn) là điều kiện chung của cả hai trường hợp, bởi chu kỳ thay thiết bị của người xem là khoảng năm tới bảy năm và không có công nghệ nào lan truyền nhanh hơn chu kỳ ấy.

> **QUY TẮC 12.7 — CÁCH ĐỌC MỘT TIN VỀ BỘ MÃ HOÁ MỚI**
>
> Cần xác định tin ấy đã qua điều kiện nào trong ba điều kiện: đã có chuẩn được ban hành chưa, đã có mạch giải mã trong thiết bị phổ thông chưa, và bản quyền sáng chế đã rõ ràng chưa.
>
> Nếu chưa qua điều kiện nào thì tin ấy là tin về nghiên cứu, không phải tin về sản xuất — và khoảng cách giữa hai loại tin ấy tính bằng nhiều năm.

12.20 Bài thực hành số 3 — kịch bản kỹ thuật cho một cảnh sản xuất ảo



Hình 12.45 Bài thực hành số 3 — đề bài, bốn phần việc và kết quả cần có

12.20.1 Đề bài

Một cảnh nội thất xe hơi đang chạy trên đường đèo, thời lượng 90 giây, quay bằng ống kính 35 mm, máy quay cách tường LED 3,5 m. Diễn viên ngồi trong xe, có kính chắn gió phản chiếu nền, và đoàn có một trường quay ảo tường cong rộng 12 m, cao 5 m.

12.20.2 Bốn phần việc

- **Chọn bước điểm ảnh.** Dùng công thức và bảng ở mục 11.24.4 với khoảng cách máy quay 3,5 m, đồng thời nêu rõ bước điểm ảnh tối đa chấp nhận được và lý do.
- **Tính điện và mạng.** Xác định số điểm ảnh của cả tường, công suất tiêu thụ ở mức sáng làm việc và lưu lượng dữ liệu cần cấp, bằng các hàm đã lập ở Chương 11.
- **Lập danh mục kiểm đồng bộ.** Danh mục gồm tần số quét tường theo góc mở cửa chập máy quay, đồng bộ giữa máy quay với tường, và độ trễ của bộ bám máy quay cùng cách đo độ trễ ấy.
- **Liệt kê việc phải xong trước ngày quay.** Theo thứ tự thời gian, ghi rõ ai làm và mất bao lâu. Đây là phần nặng nhất của bài, đúng như ngoài đời.

12.20.3 Kết quả cần có và câu hỏi quan trọng nhất

Kết quả là một trang kịch bản kỹ thuật có đủ bốn phần trên, mọi con số đều kèm phép tính và giả thiết, cùng một câu trả lời cho câu hỏi: cảnh này có nên quay ở trường quay ảo không, hay quay phông xanh với chi phí thấp hơn?

GHI NHỚ 12.5 — VÌ SAO CÂU HỎI CUỐI MỚI LÀ PHẦN QUAN TRỌNG NHẤT

Một cảnh nội thất xe có kính chắn gió phản chiếu là trường hợp tường LED có ưu thế rõ nhất: ánh sáng và phản chiếu có sẵn, đúng vật lý, và không có cách ghép nào khác cho kết quả tự nhiên bằng.

Nhưng nếu tính ra chi phí không kham nổi, thì câu trả lời "không nên" vẫn là câu trả lời đúng — miễn là có phép tính đi kèm.

Bài tập này rèn thói quen mà cả chương hướng tới: quyết định dựa trên phép tính và giả thiết viết ra được, không dựa trên cảm nhận về công nghệ mới.

12.21 Luật Trí tuệ nhân tạo của Việt Nam

12.21.1 Vì sao phần này nằm trong một cuốn sách kỹ thuật

Từ nay, quyết định dùng hay không dùng một công cụ trong sản xuất không còn thuần tuý là quyết định kỹ thuật. Một cảnh do máy sinh ra có thể hoàn hảo về hình ảnh nhưng vẫn không phát hành được nếu hồ sơ kèm theo không đầy đủ, do đó người làm nghề cần hiểu biết đủ để tự xử lý phần đơn giản và nhận biết khi nào phải hỏi người có chuyên môn pháp lý.

GHI NHỚ 12.6 — PHẠM VI CỦA PHẦN D

Phần này trình bày các nghĩa vụ ở mức đủ để áp dụng trong công việc, không ở mức phân tích pháp lý.

Mọi trường hợp có tranh chấp, có hợp đồng lớn hoặc có yếu tố nước ngoài đều phải hỏi người có chuyên môn pháp lý.

Các mốc thời gian và số hiệu văn bản trong phần này là hiện trạng tới thời điểm biên soạn; văn bản hướng dẫn thi hành vẫn đang được ban hành tiếp.

12.21.2 Văn bản và mốc thời gian

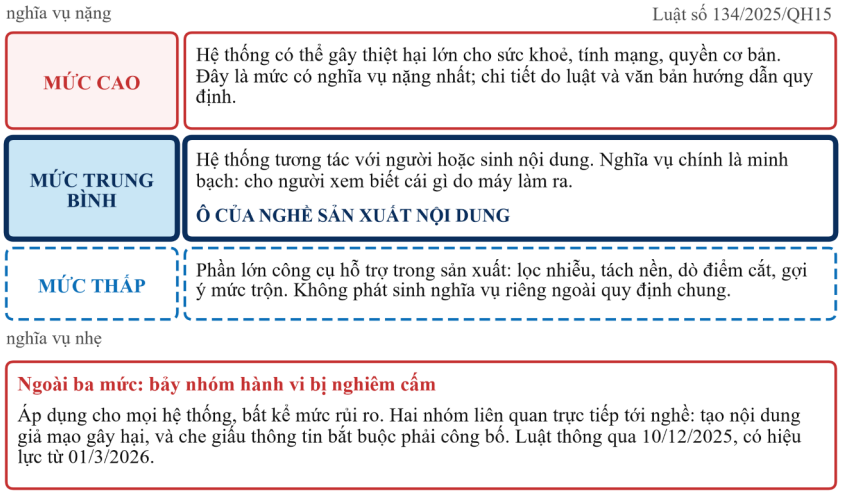
Bảng 12.8 Hai văn bản cần biết

	Việt Nam	Liên minh châu Âu
Số hiệu	Luật số 134/2025/QH15	Quy định (EU) 2024/1689
Tên gọi	Luật Trí tuệ nhân tạo	Đạo luật Trí tuệ nhân tạo của Liên minh châu Âu
Mốc thời gian	Thông qua 10/12/2025, hiệu lực 01/3/2026	Nghĩa vụ minh bạch áp dụng từ 02/8/2026
Điều về minh bạch	Điều 11	Điều 50
Điều khoản cho tác phẩm nghệ thuật	Có — gắn nhãn "theo phương thức phù hợp"	Có — miễn trừ cho tác phẩm nghệ thuật, châm biếm
Văn bản hướng dẫn	Nghị định 142/2026/NĐ-CP (ban hành 30/4/2026, hiệu lực 01/5/2026)	Gói sửa đổi "Digital Omnibus on AI": nghĩa vụ đánh dấu máy đọc được với hệ thống đã lưu hành có thể lùi tới 02/12/2026

Luật được hướng dẫn bởi Nghị định 142/2026/NĐ-CP. Việc gắn nhãn hiển thị được quy định cụ thể tại Điều 18 của nghị định này: nghĩa vụ bắt buộc áp dụng khi nội dung mô phỏng ngoại hình hoặc giọng nói của người thật, hoặc tái hiện sự kiện thực tế, và bên triển khai được chọn cách thể hiện nhãn phù hợp với loại nội dung [9]. Các mốc và điều khoản của cả hai hệ thống đều còn được cập nhật, nên người làm nghề cần tra bản đang có hiệu lực trước khi lập hồ sơ cho một dự án cụ thể [10] [1].

12.21.3 Ba mức rủi ro

Luật phân hệ thống trí tuệ nhân tạo thành ba mức rủi ro, và nghĩa vụ nặng nhẹ theo mức. Với người làm nghề sản xuất nội dung, điều quan trọng là nhận ra mình gần như luôn ở mức trung bình.



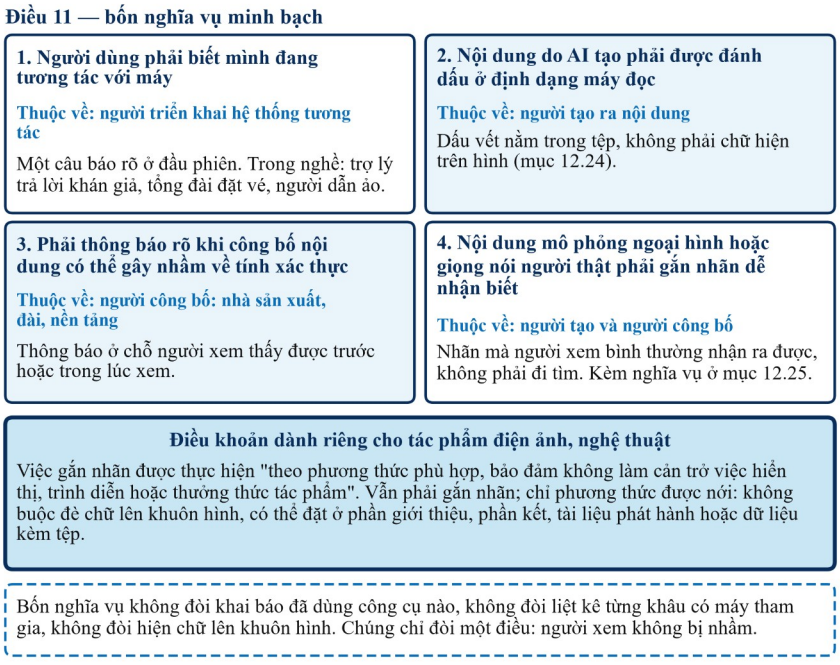
Hình 12.46 Ba mức rủi ro theo Luật số 134/2025/QH15 và ô của nghề sản xuất nội dung

Ngoài ba mức ấy, luật còn nêu bảy nhóm hành vi bị nghiêm cấm, áp dụng cho mọi hệ thống bất kể mức rủi ro. Hai nhóm liên quan trực tiếp tới nghề là tạo nội dung giả mạo gây hại và che giấu thông tin bắt buộc phải công bố; cả hai đều là những việc mà người làm nghề tử tế vốn đã không làm, nhưng nay đã có chế tài.

12.22 Điều 11 và nghĩa vụ gắn nhãn

12.22.1 Bốn nghĩa vụ

Điều 11 là điều quan trọng nhất của đạo luật đối với nghề này. Điều này đặt ra bốn nghĩa vụ không giống nhau, vì chúng thuộc về những chủ thể khác nhau và phát sinh trong những tình huống khác nhau.



Hình 12.47 Điều 11 — bốn nghĩa vụ minh bạch và điều khoản dành riêng cho tác phẩm điện ảnh, nghệ thuật

Bảng 12.9 Bốn nghĩa vụ: ai phải làm và làm gì

Nghĩa vụ	Thuộc về ai	Làm thế nào là đủ
1. Người dùng phải biết mình đang tương tác với máy	Người triển khai hệ thống tương tác	Một câu báo rõ ở đầu phiên. Không cần dài, cần rõ
2. Nội dung do AI tạo phải được đánh dấu ở định dạng máy đọc	Người tạo ra nội dung	Dấu vết nằm trong tệp, đọc được bằng máy — không phải chữ hiện trên hình. Xem mục 12.24
3. Phải thông báo rõ khi công bố nội dung có khả năng gây nhầm về tính xác thực	Người công bố: nhà sản xuất, đài, nền tảng	Thông báo ở chỗ người xem thấy được trước hoặc trong lúc xem
4. Nội dung mô phỏng ngoại hình hoặc giọng nói người thật phải gắn nhãn dễ nhận biết	Người tạo và người công bố	Nhãn mà người xem bình thường nhận ra được, không phải đi tìm mới thấy. Kèm nghĩa vụ ở mục 12.25

12.22.2 Điều mà bốn nghĩa vụ ấy KHÔNG đòi

Đây là chỗ gây lo lắng nhiều nhất và cũng là chỗ dễ làm rõ nhất. Bốn nghĩa vụ trên không đòi khai báo đã dùng công cụ nào, không đòi liệt kê từng khâu có máy tham gia, và không đòi hiện chữ lên khuôn hình.

Chúng chỉ đòi một điều: Người xem không bị nhầm. Nắm được điểm này thì phần lớn nỗi lo "phải khai báo hết mọi thứ" tự tan, và cũng thấy ngay vì sao một bộ lọc khử nhiễu không phát sinh nghĩa vụ nào: nó không làm ai nhầm về điều gì cả.

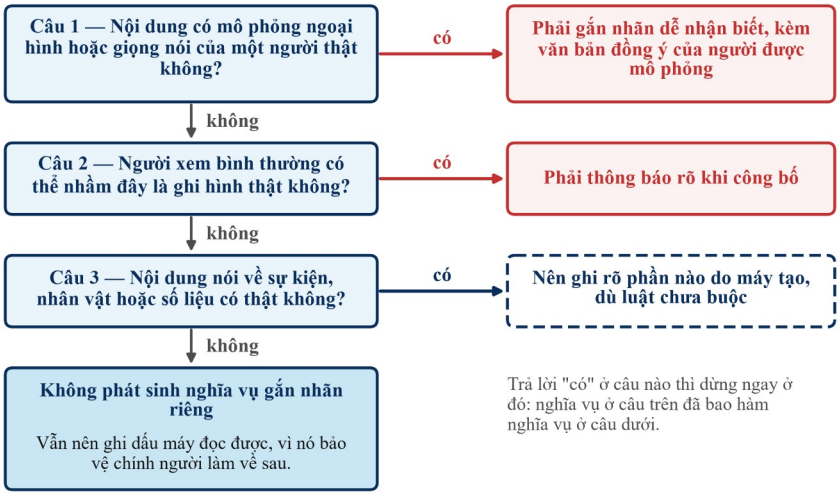
12.22.3 Điều khoản dành riêng cho tác phẩm điện ảnh, nghệ thuật

Với tác phẩm điện ảnh và nghệ thuật, luật quy định việc gắn nhãn được thực hiện "theo phương thức phù hợp, bảo đảm không làm cản trở việc hiển thị, trình diễn hoặc thưởng thức tác phẩm".

Đọc cho đúng: đây không phải miễn trừ nghĩa vụ. Vẫn phải gắn nhãn. Cái được nới là phương thức — không buộc phải đè chữ lên khuôn hình trong lúc chiếu. Nhãn có thể đặt ở phần giới thiệu đầu phim, ở phần kết, trong tài liệu phát hành, hoặc trong dữ liệu kèm tệp.

> **QUY TẮC 12.8 — CÁCH LÀM GỌN NHẤT CHO MỘT BỘ PHIM**
>
> Một dòng trong phần kết, cùng chỗ với các phần ghi công khác, nói rõ phần nào của phim có sự tham gia của công cụ tạo sinh.
>
> Kèm theo đó là dấu vết máy đọc được nằm trong tệp phát hành.
>
> Hai yếu tố này cùng nhau đáp ứng cả nghĩa vụ 2 và nghĩa vụ 3 mà không ảnh hưởng tới việc thưởng thức tác phẩm.

12.22.4 Sơ đồ quyết định



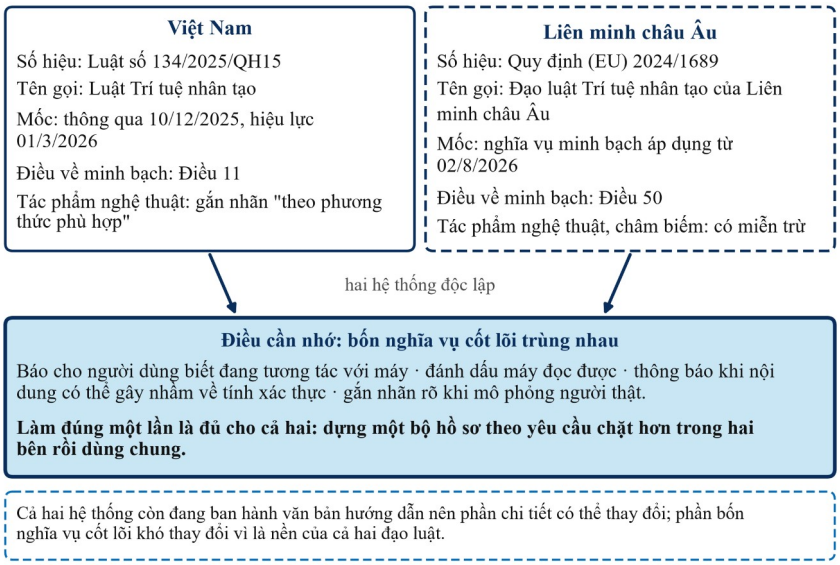
Hình 12.48 Sơ đồ ba câu hỏi: có phải gắn nhãn không, và gắn ở mức nào

Sơ đồ trên đủ dùng cho hầu hết trường hợp gặp trong sản xuất. Cần chú ý thứ tự: câu 1 có nghĩa vụ nặng nhất, và khi trả lời "có" ở câu nào thì dừng ngay ở đó, không cần xét tiếp, vì nghĩa vụ ở câu trên đã bao hàm nghĩa vụ ở câu dưới.

12.23 Nghĩa vụ tương ứng ở thị trường ngoài nước

12.23.1 Vì sao mục này cần

Người làm nghề ngày nay thường làm phim, phim quảng cáo và nội dung mạng bán ra nước ngoài, hoặc gia công cho đối tác nước ngoài; nếu chỉ biết quy định trong nước thì sẽ phải làm lại hồ sơ, hoặc nặng hơn là bị từ chối nhận hàng [2].



Hình 12.49 Hai hệ thống pháp luật độc lập với yêu cầu gần như giống nhau

12.23.2 Điều quan trọng nhất: bốn nghĩa vụ trùng nhau

Đạo luật Trí tuệ nhân tạo của Liên minh châu Âu, Điều 50, áp dụng từ 02/8/2026 (riêng nghĩa vụ đánh dấu máy đọc được với hệ thống đã lưu hành có thể

lùi tới 02/12/2026 theo gói sửa đổi của Liên minh), đặt ra bốn nghĩa vụ gần như song song với Điều 11 của luật Việt Nam — kể cả điều khoản nới phương thức gắn nhãn cho tác phẩm nghệ thuật và châm biếm.

> **GHI NHỚ 12.7 — HỆ QUẢ THỰC HÀNH, VÀ NÓ LÀ HỆ QUẢ TỐT**
> Làm đúng một lần là đủ cho cả hai. Không cần hai quy trình, không cần hai bộ hồ sơ. Cách làm là dựng một bộ hồ sơ theo yêu cầu chặt hơn trong hai bên rồi dùng chung. Điều này không phải ngẫu nhiên, vì hai đạo luật cùng xuất phát từ một nhận định về điều cần bảo vệ: người xem không bị nhầm, và người bị mô phỏng không bị dùng hình ảnh khi chưa đồng ý.

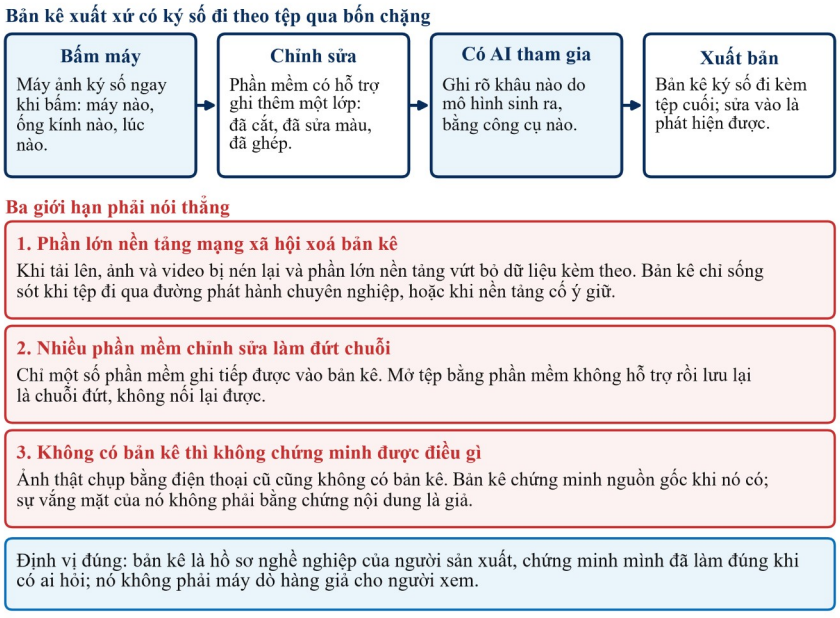
Cần lưu ý về thời gian: hai hệ thống có lịch áp dụng khác nhau và đều còn đang ban hành văn bản hướng dẫn, nên phần chi tiết có thể thay đổi. Phần bốn nghĩa vụ cốt lõi khó thay đổi vì là nền của cả hai đạo luật, và đó là lý do mục này được viết theo nghĩa vụ chứ không theo điều khoản.

Với thị trường ngoài hai hệ thống trên, nguyên tắc thực hành vẫn giữ nguyên: tra xem nơi phát hành có quy định riêng không, và nếu không tra được thì làm theo bộ hồ sơ đã dựng. Một bộ hồ sơ đáp ứng được yêu cầu chặt hơn trong hai hệ thống đã bao phủ phần lớn những gì các nơi khác đòi hỏi, vì phần lớn các nơi ấy cũng đang lấy hai hệ thống này làm mẫu.

12.24 Xuất xứ nội dung: công cụ và giới hạn thật

12.24.1 Đây là công cụ để thực hiện nghĩa vụ thứ hai

Nghĩa vụ "đánh dấu ở định dạng máy đọc" cần một cách làm cụ thể. Cách làm đang được ngành công nghiệp thống nhất là gắn vào tệp một bản kê xuất xứ có ký số: một khối dữ liệu ghi lại nguồn gốc của tệp và các bước chỉnh sửa đã trải qua, ký bằng chữ ký số nên sửa vào là phát hiện được [6].



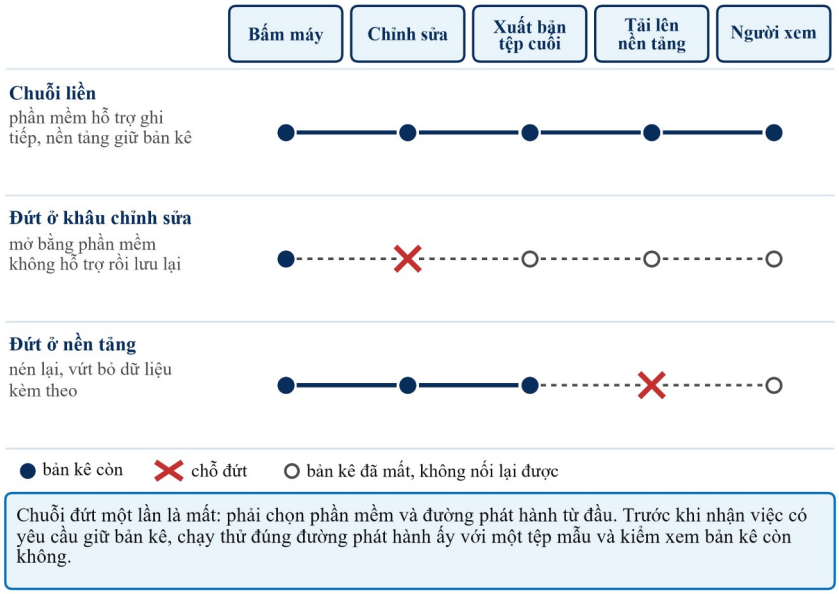
Hình 12.50 Bản kê xuất xứ đi theo tệp qua bốn chặng, và ba giới hạn phải nói thẳng

Tới thời điểm biên soạn, một số dòng máy ảnh chuyên nghiệp đã ký số ngay khi bấm máy (máy nào, ống kính nào, lúc nào), và một số phần mềm dựng và sửa ảnh ghi tiếp được vào bản kê ấy. Về nguyên tắc, một tệp có thể mang theo lý lịch đầy đủ từ lúc bấm máy tới lúc phát hành.

12.24.2 Ba giới hạn phải nói thẳng với người đọc

Đó là về nguyên tắc; trên thực tế có ba giới hạn mà người đọc cần được biết rõ.

- **Phần lớn nền tảng mạng xã hội xoá bản kê.** Khi tải lên, ảnh và video bị nén lại theo cách của nền tảng; phần lớn nền tảng vứt bỏ dữ liệu kèm theo. Bản kê chỉ sống sót khi tệp đi qua đường phát hành chuyên nghiệp, hoặc khi nền tảng cố ý giữ.
- **Nhiều phần mềm chỉnh sửa làm đứt chuỗi.** Chỉ một số phần mềm ghi tiếp được vào bản kê. Mở tệp bằng một phần mềm không hỗ trợ rồi lưu lại là chuỗi đứt, và không nối lại được.
- **Không có bản kê thì không chứng minh được điều gì.** Một bức ảnh thật chụp bằng điện thoại cũ cũng không có bản kê. Bản kê chứng minh được nguồn gốc khi nó có; sự vắng mặt của nó không phải bằng chứng cho thấy nội dung là giả.



Hình 12.51 Chuỗi xuất xứ đứt ở đâu — hai chỗ đứt phổ biến nhất

GHI NHỚ 12.8 — ĐỊNH VỊ ĐÚNG CÔNG CỤ NÀY

Ba giới hạn trên không làm bản kê xuất xứ trở nên vô dụng mà chỉ xác định đúng phạm vi sử dụng của nó.

Bản kê là hồ sơ nghề nghiệp của người sản xuất — thứ chứng minh mình đã làm đúng khi có ai hỏi. Nó không phải máy dò hàng giả cho người xem.

Về thực hành, trước khi nhận một việc có yêu cầu giữ bản kê, cần chạy thử đúng đường phát hành ấy với một tệp mẫu và kiểm tra bản kê còn hay không, vì kiểm sau khi làm xong thì đã muộn.

12.25 Quyền hình ảnh và giọng nói — bản sao số của người biểu diễn

12.25.1 Hai loại bản sao số

Cần phân biệt hai trường hợp vì nghĩa vụ khác nhau [4]:

- **Bản sao số tạo ra trong quá trình làm việc.** Diễn viên đã đến trường quay, đã diễn, và đoàn ghi thêm dữ liệu để dựng bản sao số phục vụ chính dự án ấy — chẳng hạn để quay cảnh nguy hiểm, hoặc để sửa một câu thoại.
- **Bản sao số tạo ra độc lập.** Dựng từ tư liệu đã có sẵn, không cần người ấy tham gia, và dùng cho việc mà người ấy chưa biết. Đây là trường hợp nhiều rủi ro hơn hẳn.

12.25.2 Sáu điểm bắt buộc của một văn bản đồng ý



Hình 12.52 Sáu điểm bắt buộc của một văn bản đồng ý cho bản sao số

Sáu điểm dưới đây lấy từ thoả ước nghề nghiệp đang có hiệu lực ở thị trường nước ngoài, nơi vấn đề đã được thương lượng kỹ giữa người sử dụng lao động và tổ chức của người biểu diễn. Pháp luật trong nước chưa quy định chi tiết tới mức này, nhưng một văn bản đủ sáu điểm thì an toàn ở cả hai nơi [8].

Bảng 12.10 Sáu điểm bắt buộc và lỗi thường gặp

Điểm	Yêu cầu	Lỗi thường gặp
1. Tách riêng và ký riêng	Một văn bản riêng, chữ ký riêng, phần cốt lõi in đậm và viết hoa	Giấu thành một điều khoản trong hợp đồng chung
2. Mô tả cụ thể mục đích	Dùng vào cảnh nào, làm gì, trong sản phẩm nào	"Sử dụng cho các mục đích quảng bá" — không đủ cụ thể
3. Giới hạn theo từng dự án	Chỉ cho dự án đang làm; dự án sau phải ký lại	Đồng ý trọn gói cho mọi dự án tương lai — điều khoản bị lạm dụng nhiều nhất
4. Thù lao	Ít nhất bằng mức ngày công tối thiểu cho thời gian ghi dữ liệu, cộng khoản trả cho từng lần sử dụng lại	Coi việc ghi dữ liệu là phần của ngày quay, không trả thêm
5. Quyền sau khi dự án kết thúc	Dữ liệu đã ghi được giữ hay xoá, ai giữ, giữ bao lâu, có được dùng để huấn luyện mô hình khác không	Không nhắc tới — mặc định thành giữ vô thời hạn
6. Hiệu lực sau khi qua đời	Chỉ có hiệu lực nếu người thừa kế hoặc người đại diện hợp pháp đồng ý	Coi như văn bản cũ vẫn còn hiệu lực

12.25.3 Mẫu văn bản rút gọn

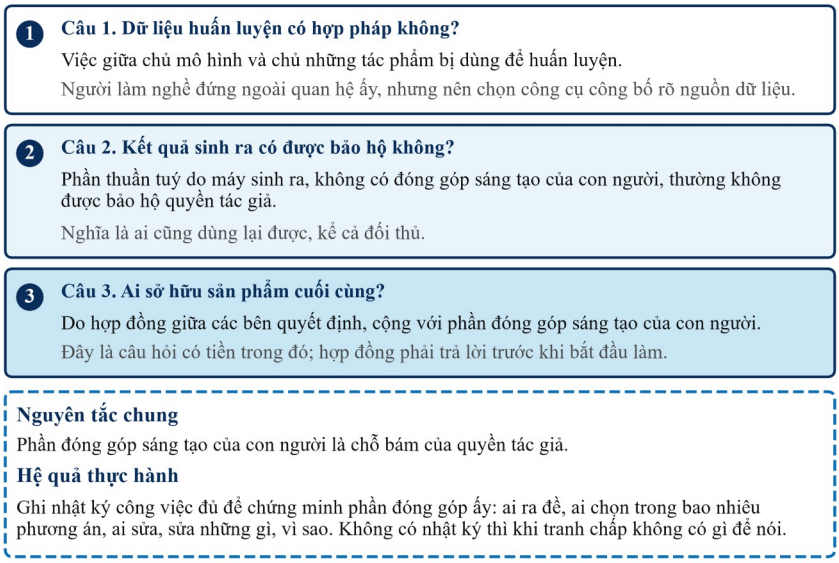
Dưới đây là khung một văn bản đồng ý rút gọn, đủ sáu điểm, dùng cho bài tập và cho các dự án nhỏ; với dự án có giá trị lớn, văn bản phải do người có chuyên môn pháp lý soạn.

- Phần 1 — Người đồng ý: họ tên, ngày sinh, giấy tờ tuỳ thân và vai trò trong dự án.
- Phần 2 — Đối tượng: mô tả rõ dữ liệu sẽ được ghi (hình ảnh khuôn mặt, dáng người hoặc bản thu giọng, nêu cụ thể loại nào), nơi ghi và thời gian ghi.
- Phần 3 — Mục đích sử dụng: tên dự án, các cảnh cụ thể, cách sử dụng cụ thể. Có câu khẳng định văn bản này không áp dụng cho bất kỳ dự án nào khác.
- Phần 4 — Thù lao: khoản trả cho thời gian ghi dữ liệu và khoản trả cho từng lần sử dụng lại nếu có.
- Phần 5 — Sau khi dự án kết thúc: bên nào giữ dữ liệu, giữ tới bao giờ, xoá bằng cách nào, có được dùng để huấn luyện mô hình không (mặc định là không).
- Phần 6 — Hiệu lực và chấm dứt: thời hạn, quyền rút lại và quy định về hiệu lực sau khi người đồng ý qua đời.
- Phần 7 — Chữ ký riêng của người đồng ý cho chính văn bản này, kèm ngày.

12.26 Quyền tác giả đối với nội dung có AI tham gia

12.26.1 Ba câu hỏi tách bạch

Nhiều người thường gộp ba câu hỏi khác nhau làm một rồi đi tới kết luận sai, trong khi nếu tách ra thì cả ba câu đều dễ trả lời hơn nhiều [3].



Hình 12.53 Ba câu hỏi quyền tác giả mà nhiều người hay gộp làm một

- **Câu 1 — dữ liệu huấn luyện có hợp pháp không?** Đây là vấn đề giữa chủ mô hình và chủ những tác phẩm bị dùng để huấn luyện, và người làm nghề đứng ngoài quan hệ ấy. Tuy nhiên, nên chọn công cụ có công bố rõ nguồn dữ liệu, vì rủi ro của chủ mô hình cuối cùng cũng ảnh hưởng tới người dùng.
- **Câu 2 — kết quả sinh ra có được bảo hộ không?** Phần thuần tuý do máy sinh ra, không có đóng góp sáng tạo của con người, thường không được bảo hộ quyền tác giả. Nghĩa là ai cũng dùng lại được — kể cả đối thủ.
- **Câu 3 — ai sở hữu sản phẩm cuối cùng?** Quyền sở hữu do hợp đồng giữa các bên quyết định, cộng với phần đóng góp sáng tạo của con người. Đây là câu hỏi có ý nghĩa kinh tế lớn nhất, và một hợp đồng chặt chẽ phải trả lời câu hỏi này trước khi bắt đầu làm.

12.26.2 Nguyên tắc chung và hệ quả thực hành

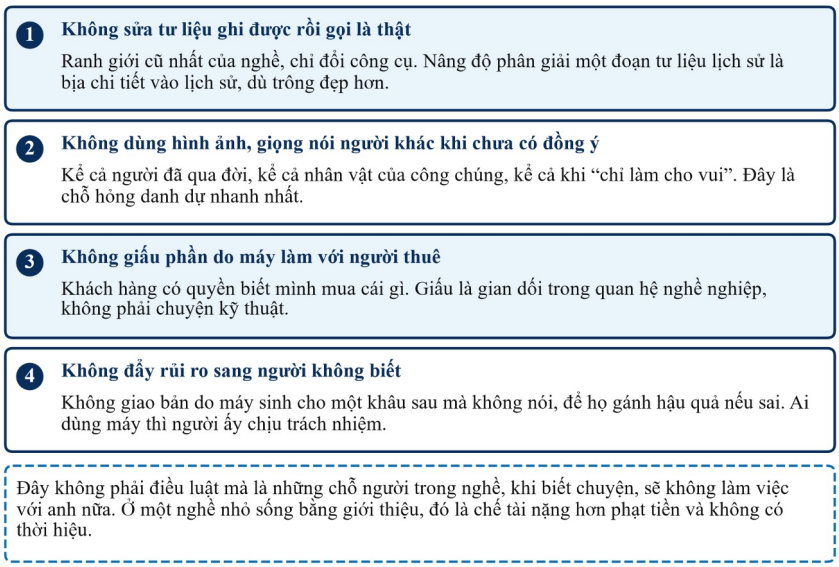
> **QUY TẮC 12.9 — NGUYÊN TẮC BÁM VÀO ĐƯỢC**
>
> Phần đóng góp sáng tạo của con người là chỗ bám của quyền tác giả.
>
> Hệ quả là phải ghi nhật ký công việc đủ để chứng minh phần đóng góp ấy: ai ra đề, ai chọn trong bao nhiêu phương án, ai sửa, sửa những gì và vì sao.
>
> Nhật ký nghe như một thủ tục phiền hà, nhưng đó là căn cứ duy nhất phân biệt được việc "tôi làm ra tác phẩm này" với việc "tôi bấm nút bốn lần"; khi có tranh chấp, thiếu nhật ký thì không còn gì để chứng minh.

Cách ghi nhật ký gọn nhất mà vẫn đủ là một tệp bảng tính, mỗi dòng ứng với một cảnh, các cột gồm: lời đề bài đã dùng, số phương án đã sinh, phương án được chọn và lý do, các bước sửa tay sau đó, người thực hiện và ngày thực hiện.

Nhật ký cần được ghi ngay trong lúc làm, vì ghi bù sau khi xong dự án vừa mất thời gian vừa kém tin cậy.

12.27 Đạo đức nghề và cách đọc một dự báo công nghệ

12.27.1 Bốn ranh giới không nên vượt dù luật chưa cấm

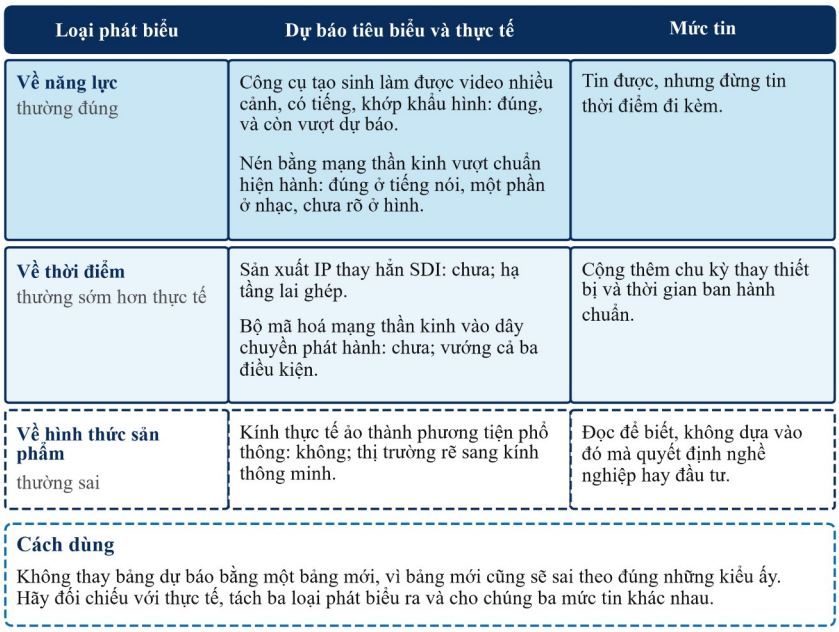


Hình 12.54 Bốn ranh giới nghề nghiệp không nên vượt, dù luật chưa cấm

Bốn ranh giới ở hình trên không có hiệu lực pháp lý, nhưng người trong nghề, khi biết chuyện, sẽ ngừng hợp tác với người vi phạm; ở một nghề nhỏ và sống nhờ giới thiệu, đó là chế tài nặng hơn phạt tiền và không có thời hiệu.

12.27.2 Đối chiếu một bảng dự báo với thực tế

Một bảng dự báo công nghệ tới năm 2030 nên được dùng theo cách này: không thay nó bằng một bảng dự báo mới — vốn cũng sẽ sai theo đúng những kiểu như thế — mà đối chiếu nó với thực tế đã xảy ra và rút ra quy luật.



Hình 12.55 Đối chiếu bảng dự báo với thực tế và ba quy luật rút ra được

Bảng 12.11 Đối chiếu một số dự báo tiêu biểu với thực tế

Dự báo tiêu biểu	Thực tế tới thời điểm biên soạn	Loại
Công cụ tạo sinh sẽ làm ra được video nhiều cảnh, có tiếng, khớp khẩu hình	Đúng. Năng lực này đã có và còn vượt mức dự báo	Năng lực — đúng
Nén dùng mạng thần kinh sẽ vượt các chuẩn hiện hành	Đúng ở tiếng nói, đúng một phần ở nhạc, chưa rõ ở hình	Năng lực — đúng một phần
Sản xuất trên nền IP sẽ thay hẳn SDI	Chưa. Hạ tầng lai ghép, và SDI vẫn giữ vững ở đầu cuối và đơn vị nhỏ	Thời điểm — sớm
Bộ mã hoá dùng mạng thần kinh sẽ vào dây chuyền phát hành	Chưa. Vướng cả ba điều kiện ở mục 12.19.3	Thời điểm — sớm
Kính thực tế ảo sẽ trở thành phương tiện phổ thông	Không. Thị trường rẽ sang kính thông minh; sản phẩm cao cấp được nêu tên đã dừng sản xuất	Hình thức sản phẩm — sai

12.27.3 Ba quy luật, phát biểu ở dạng dùng lại được

- **Dự báo về năng lực thường đúng.** Vì năng lực là chuyện kỹ thuật, mà kỹ thuật thì đi theo đường đã nhìn thấy trước.
- **Dự báo về thời điểm thường sớm hơn thực tế.** Vì người dự báo tính theo tiến bộ kỹ thuật, quên chu kỳ thay thiết bị của người dùng và thời gian ban hành chuẩn.
- **Dự báo về hình thức sản phẩm thường sai.** Vì hình thức do người dùng chọn, mà người dùng không chọn theo thông số kỹ thuật.

QUY TẮC 12.10 — CÁCH DÙNG BA QUY LUẬT NÀY

Khi đọc bất kỳ dự báo công nghệ nào, kể cả dự báo trong chương này, người đọc cần tách riêng ba loại phát biểu ấy và gán cho chúng ba mức tin cậy khác nhau.

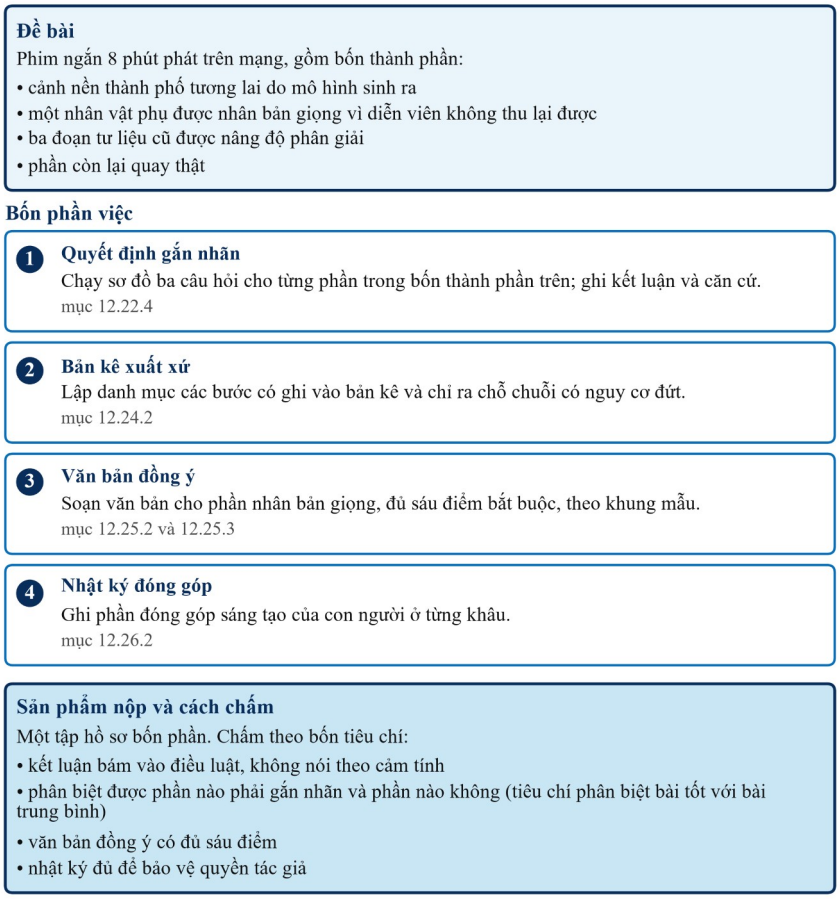
Phát biểu về năng lực có thể tin được, nhưng không nên tin theo mốc thời điểm đi kèm.

Phát biểu về hình thức sản phẩm sẽ thắng chỉ nên đọc để biết, không nên làm căn cứ cho quyết định nghề nghiệp hay đầu tư.

Một cách dùng nữa: một bảng dự báo như thế nên được giữ lại trong mọi lần tái bản, kèm cột đối chiếu cập nhật. Sau mỗi vài năm, chính bảng ấy lại dạy được một bài học mới mà không cần viết thêm dòng nào — vì thực tế đã viết hộ. Rất ít nội dung trong một cuốn sách kỹ thuật có được tính chất ấy: càng cũ đi càng dạy được nhiều hơn.

12.28 Bài thực hành số 4, bảng tra nhanh và tổng kết

12.28.1 Bài thực hành số 4 — hồ sơ pháp lý cho một sản phẩm có phần do AI tạo



Hình 12.56 Bài thực hành số 4 — đề bài, bốn phần việc và cách tự đánh giá

Đề bài: một phim ngắn 8 phút phát trên mạng, trong đó cảnh nền thành phố tương lai do mô hình sinh ra, một nhân vật phụ được nhân bản giọng vì diễn viên

không thể thu lại, ba đoạn tư liệu cũ được nâng độ phân giải, và phần còn lại quay thật.

- **Phần 1 — quyết định gắn nhãn.** Chạy sơ đồ ba câu hỏi ở mục 12.22.4 cho từng phần trong bốn phần trên. Ghi kết luận và căn cứ.
- **Phần 2 — bản kê xuất xứ.** Lập danh mục các bước có ghi vào bản kê và chỉ ra chỗ chuỗi có nguy cơ đứt theo mục 12.24.2.
- **Phần 3 — văn bản đồng ý.** Soạn văn bản cho phần nhân bản giọng, đủ sáu điểm ở mục 12.25.2, theo khung mẫu ở mục 12.25.3.
- **Phần 4 — nhật ký đóng góp.** Ghi phần đóng góp sáng tạo của con người ở từng khâu, theo cách ghi ở mục 12.26.2.

THỰC HÀNH 12.3 — CÁCH TỰ ĐÁNH GIÁ BÀI 4

Kết luận có bám vào điều luật hay chỉ dựa trên cảm tính.

Có phân biệt được phần nào phải gắn nhãn và phần nào không; đây là điểm phân biệt bài tốt với bài trung bình.

Văn bản đồng ý có đủ sáu điểm không.

Nhật ký có đủ để bảo vệ quyền tác giả không.

Bài này không đòi hỏi kiến thức luật chuyên sâu mà đòi hỏi đọc kỹ một điều luật ngắn và áp dụng đúng vào bốn tình huống cụ thể.

12.28.2 Bảng tra nhanh

Tình huống	Làm gì	Nghĩa vụ
Dùng AI lọc nhiễu, tách nền, dò điểm cắt	Không phát sinh nghĩa vụ. Ghi vào nhật ký công việc là đủ.	Không
Nâng độ phân giải tư liệu để phát hành	Phải ghi rõ đã nâng cấp. Với tư liệu tài liệu, lịch sử: bắt buộc.	Ghi nhận
Sinh cảnh nền, vật thể phụ bằng mô hình	Đánh dấu máy đọc được trong tệp. Không buộc hiện chữ lên hình.	Ghi nhận
Nhân bản giọng của diễn viên trong đoàn	Văn bản đồng ý đủ sáu điểm, cộng đánh dấu máy đọc được.	Nặng
Mô phỏng ngoại hình hoặc giọng người thật	Văn bản đồng ý, cộng nhãn dễ nhận biết cho người xem. Nặng nhất.	Nặng nhất
Nội dung có thể bị nhầm là ghi hình thật	Phải thông báo rõ khi công bố. Nghĩa vụ của người công bố.	Nặng

Căn cứ

Luật số 134/2025/QH15, Điều 11, hiệu lực từ 01/3/2026. Thị trường ngoài nước: Quy định (EU) 2024/1689, Điều 50, áp dụng từ 02/8/2026.

Hình 12.57 Bảng tra nhanh — làm gì trong từng tình huống

Bảng trên là bản rút gọn để dán lên tường phòng dựng và không thay thế được việc đọc điều luật khi gặp trường hợp khó. Trước khi in ra dùng chính thức, bảng cần được người có chuyên môn pháp lý soát lại và cập nhật khi có văn bản hướng dẫn thi hành mới.

12.28.3 Tổng kết Chương 12

Chương này khác mười một chương trước ở chỗ nội dung của nó có khả năng lạc hậu nhanh, nên cấu trúc được dựng để chịu được điều đó: nguyên lý và khung đánh giá ở phần thân, danh mục công cụ ở phụ lục có thể thay hằng năm.

- **Phần A** trình bày hai trục phân loại và hai phép đo: ba loại tác vụ, ba mức can thiệp, tỉ lệ kết quả dùng được và thời gian thật. Đây là phần có tuổi thọ dài nhất của chương.
- **Phần B** đi qua từng khâu của dây chuyền và trả lời câu hỏi nên dùng hay không bằng chi phí thật và điểm hoà vốn, thay cho cảm giác.
- **Phần C** nối phần cứng của Chương 11 với quy trình, gồm sản xuất ảo, hạ tầng lai ghép và cách đọc một phương tiện mới mà không cần bảng thiết bị.
- **Phần D** trình bày bốn nghĩa vụ minh bạch, một sơ đồ quyết định gắn nhãn, sáu điểm của văn bản đồng ý và ba câu hỏi về quyền tác giả.

12.28.4 Tổng kết cả cuốn sách

Mười hai chương đã đi hết một vòng: từ bản chất vật lý của âm thanh và ánh sáng, qua việc biến chúng thành con số, qua nén và lưu trữ, qua ghi hình và ghi tiếng, qua hậu kỳ, rồi trở lại thành ánh sáng và âm thanh cho người xem, và cuối cùng là những công cụ mới đang xen vào toàn bộ vòng ấy.

> **GHI NHỚ 12.9 — BA ĐIỀU MÀ CẢ CUỐN SÁCH MUỐN ĐỂ LẠI**
>
> Thứ nhất — mọi con số trong nghề đều phải truy được về một phép đo hoặc một phép tính. Con số không truy được là con số đi mượn, và đi mượn thì có ngày phải trả.
>
> Thứ hai — công cụ đổi liên tục, nguyên lý thì không. Người nắm nguyên lý sẽ dùng được công cụ chưa ra đời; người đọc thuộc công cụ sẽ phải học lại sau mỗi ba năm.
>
> Thứ ba — mọi quyết định kỹ thuật cuối cùng đều là quyết định của một con người, và người ấy chịu trách nhiệm. Máy chuẩn bị được mọi thứ cho quyết định ấy, nhưng chữ ký thì phải là chữ ký của người.

Điều cuối cùng dành riêng cho chương này: nếu năm năm sau đọc lại mà thấy Phần B và Phần C đã lạc hậu thì đó là điều bình thường và đã được lường trước; còn nếu Phần A cũng lạc hậu, tức là ba loại tác vụ, ba mức can thiệp và cách đo tỉ lệ kết quả dùng được không còn dùng được nữa, thì chương này đã chọn sai vật liệu và người biên soạn bản sau nên bắt đầu lại từ chỗ ấy.

TÓM TẮT CHƯƠNG 12

1. Công cụ AI được phân theo ba loại tác vụ (nhận dạng, biến đổi, tạo sinh) và ba mức can thiệp; hai trục này độc lập và bền hơn mọi tên sản phẩm.
2. Hai số đo của một công cụ tạo sinh là tỉ lệ kết quả dùng được η và thời gian thật cho một kết quả dùng được; công cụ được đánh giá bằng sáu câu hỏi về chỗ đứng của nó trong công việc thật.
3. Tiền kỳ, trường quay, hậu kỳ hình và hậu kỳ tiếng có mức chín muồi khác nhau: dọn tiếng và nhận dạng chín muồi nhất, còn tạo sinh hình và giọng nhân bản có nhiều rủi ro nhất.

4. Quyết định dùng hay không dựa vào chi phí thật và điểm hoà vốn, không dựa vào cảm giác.
5. Sản xuất ảo nối tường LED của Chương 11 với quy trình mới, trong đó khối lượng công việc dời sang tiền kỳ, kỹ xảo hoàn thành ngay trong máy quay và ê-kíp có thêm ba vai trò mới.
6. Hạ tầng sản xuất hiện là lai ghép giữa IP, SDI và đám mây, và bài toán của dựng từ xa là ngân sách độ trễ; nén bằng mạng thần kinh vượt trội rõ rệt ở tiếng nói nhưng chưa vào dây chuyền phát hành.
7. Nghĩa vụ minh bạch, gắn nhãn, xuất xứ nội dung, văn bản đồng ý về bản sao số và quyền tác giả là phần việc kỹ thuật của người làm nghề, không chỉ thuộc về luật sư.
8. Dự báo về năng lực thường đúng, về thời điểm thường sớm, về hình thức sản phẩm thường sai; người đọc nên tách ba loại phát biểu này khi đọc mọi dự báo công nghệ.

12.29 Câu hỏi và bài tập ôn tập

A. Câu hỏi lý thuyết

Phần A — nguyên lý và khung đánh giá

1. Vì sao một chương về công nghệ mới lại phải bắt đầu bằng việc bàn về vật liệu viết? Nêu ba bậc tuổi thọ của vật liệu và cho ví dụ mỗi bậc.
2. Phân biệt ba loại tác vụ: nhận dạng, biến đổi và tạo sinh, và cho biết quan hệ giữa đầu vào và đầu ra của mỗi loại.
3. Vì sao nhóm tác vụ biến đổi là nhóm cần cẩn thận nhất, dù nó không sinh ra nội dung mới?
4. Phát biểu quy tắc phân biệt giữa nâng cấp tư liệu hư cấu và nâng cấp tư liệu tài liệu, và giải thích vì sao ranh giới ấy quan trọng.
5. Tính không lặp lại của tác vụ tạo sinh dẫn tới hệ quả thực hành gì trong việc lưu trữ?
6. Nêu ba mức can thiệp của máy vào một khâu công việc. Điều gì thay đổi theo từng mức, và điều gì không thay đổi?
7. Vì sao máy xen vào mạnh nhất ở hai đầu của dây chuyền mà không xen vào khâu giữa, và lý do ấy có thuộc về kỹ thuật hay không?
8. Nêu ba mốc bàn giao không giao được cho máy và giải thích vì sao; giới hạn ấy thuộc về năng lực hay thuộc về trách nhiệm?
9. Định nghĩa tỉ lệ kết quả dùng được và giải thích vì sao nó không phải là một thuộc tính của riêng công cụ.
10. Viết công thức thời gian thực tế cho một kết quả dùng được và giải thích vì sao phải chia cho η.
11. Trong ví dụ ở mục 12.5.2, phần lớn thời gian nằm ở khâu nào, và hệ quả đối với việc muốn làm nhanh hơn là gì?

12. Mô tả cách tự đo η trên mười việc điển hình. Điều gì phải làm trước khi sinh kết quả đầu tiên, và vì sao?
13. Nêu sáu câu hỏi để đánh giá một công cụ; câu nào thường bị bỏ qua nhất và bỏ qua nó dẫn tới rủi ro gì?

Phần B — AI ở từng khâu của dây chuyền

1. Ở tiền kỳ, điều mà máy làm tốt hơn người không phải là chất lượng mà là gì? Giải thích bằng ví dụ phân cảnh bằng hình.
2. Bẫy trung bình là gì, vì sao nó nguy hiểm một cách âm thầm và phòng tránh bằng cách nào?
3. Vì sao khâu trường quay là khâu ít gây tranh cãi nhất trong cả dây chuyền?
4. Bám nét tự động tối ưu điều gì, vì sao điều đó có thể khác với ý đồ của đạo diễn, và quy tắc sử dụng là gì?
5. Vì sao nhóm nhận dạng và biến đổi ở hậu kỳ hình là nhóm chín muồi nhất? Nêu hai tính chất làm nên điều đó.
6. Khi một công đoạn rút từ ba ngày xuống ba giờ, ngoài chi phí thì điều gì thay đổi, và điều đó có thể gây hậu quả gì?
7. Nêu ba điều mà công cụ tạo sinh hình chưa làm được, và giải thích vì sao đó chính là ba điều một bộ phim cần nhất.
8. Vị trí đúng của công cụ tạo sinh hình hiện nay là công cụ làm phim hay công cụ của khâu kỹ xảo? Lập luận.
9. Vì sao công cụ dùng mạng thần kinh vượt trội hoàn toàn ở dọn tiếng? Giải thích bằng bản chất của tín hiệu tiếng nói.
10. Vì sao nhóm dọn tiếng gần như không còn tranh cãi về mặt pháp lý?
11. Vì sao nhóm giọng nhân bản là nhóm nhiều rủi ro nhất của cả chương, và quy tắc tối thiểu phải theo là gì?
12. Viết công thức chi phí thật cho một kết quả dùng được và công thức điểm hoà vốn, đồng thời nêu điều kiện tồn tại của điểm hoà vốn.
13. Trong bảng ba giá trị của η ở mục 12.14.2, vì sao cùng một công cụ lại cho ba kết luận khác nhau?
14. Vì sao con số dung lượng phải chuẩn bị chỗ là con số cao điểm chứ không phải con số sau khi lọc?

Phần C: sản xuất ảo, phương tiện mới và hạ tầng

1. Nêu ba nguồn đầu vào của một máy dựng cảnh thời gian thực và nhiệm vụ của từng nguồn.
2. Phân biệt vùng nhìn của máy quay với phần tường còn lại, và giải thích vì sao cách chia ấy tiết kiệm được nhiều sức máy.
3. Nếu bộ bám máy quay bị trễ thì hiện tượng gì xảy ra trên khuôn hình, và vai trò nào trong ê-kíp chịu trách nhiệm về việc này?
4. Ở 48 hình một giây, máy dựng có bao nhiêu thời gian cho một khung hình, và vì sao đây là ràng buộc cứng?

5. Nêu bốn lợi ích của ICVFX; lợi ích nào trong bốn lợi ích ấy không thể thay thế bằng bất kỳ kỹ thuật hậu kỳ nào?
6. Cái giá ít được nói trước của sản xuất ảo là gì? Vẽ lại hình dáng của hai biểu đồ khối lượng công việc và giải thích hệ quả quản lý.
7. Nêu ba dấu hiệu cho thấy một cảnh không nên quay ở trường quay ảo.
8. Nêu ba vai trò mới trong ê-kíp sản xuất ảo và công việc chính của từng vai trò.
9. Vì sao mục về XR không có bảng liệt kê thiết bị, và nội dung thay thế là gì?
10. Trình bày câu chuyện rẽ hướng của thị trường thiết bị đội đầu và rút ra bài học về ba loại phát biểu trong một dự báo.
11. Nêu sáu câu hỏi để đọc một phương tiện mới; câu nào nặng nhất và vì sao?
12. Vì sao dự báo "IP sẽ thay hẳn SDI" chỉ đúng một nửa, và ba điều kiện để một đơn vị chuyển sang IP là gì?
13. Viết công thức độ trễ vòng khi dựng từ xa, nêu ba ngưỡng cảm nhận và ý nghĩa của từng ngưỡng.
14. Vì sao canh màu cuối và trộn tiếng phòng lớn vẫn chưa đưa lên đám mây được, và lý do ấy có phải là băng thông hay không?
15. Vì sao bộ mã hoá dùng mạng thần kinh vượt trội rõ rệt ở tiếng nói mà chỉ hơn ít ở nhạc và ở hình?
16. Nêu ba điều kiện để một bộ mã hoá mới thật sự được dùng; điều kiện nào liên quan tới chu kỳ thay thiết bị của người xem?

Phần D: pháp lý, xuất xứ nội dung và đạo đức nghề

1. Nêu số hiệu, ngày thông qua và ngày có hiệu lực của luật về trí tuệ nhân tạo hiện hành ở Việt Nam.
2. Nêu ba mức rủi ro của hệ thống trí tuệ nhân tạo; nghề sản xuất nội dung thường nằm ở mức nào và vì sao?
3. Nêu bốn nghĩa vụ minh bạch và cho biết mỗi nghĩa vụ thuộc về ai.
4. Bốn nghĩa vụ ấy không đòi những gì? Vì sao điểm này quan trọng?
5. Điều khoản dành riêng cho tác phẩm điện ảnh, nghệ thuật nới lỏng những nghĩa vụ nào và giữ nguyên những nghĩa vụ nào?
6. Vẽ lại sơ đồ ba câu hỏi quyết định gắn nhãn, và giải thích vì sao khi câu trả lời là "có" ở một câu thì việc xét dừng ngay tại câu đó.
7. Vì sao người làm nghề ở Việt Nam cần biết nghĩa vụ tương ứng ở thị trường ngoài nước, và hệ quả thực hành của điều đó là gì?
8. Bản kê xuất xứ nội dung là công cụ để thực hiện nghĩa vụ nào, và nguyên lý hoạt động của nó ra sao?
9. Nêu ba giới hạn thực tế của bản kê xuất xứ, và giải thích vì sao việc không có bản kê chưa phải là bằng chứng.
10. Chuỗi xuất xứ thường đứt ở hai khâu nào, cách phòng ngừa là gì, và vì sao phải phòng ngừa ngay từ đầu?

11. Phân biệt bản sao số tạo ra trong quá trình làm việc với bản sao số tạo ra độc lập.
12. Nêu sáu điểm bắt buộc của một văn bản đồng ý cho bản sao số, và cho biết điểm nào bị lạm dụng nhiều nhất.
13. Tách ba câu hỏi về quyền tác giả mà nhiều người thường gộp làm một, và cho biết câu nào gắn với lợi ích kinh tế.
14. Nguyên tắc chung của quyền tác giả đối với nội dung có AI tham gia là gì, và hệ quả thực hành đối với việc ghi chép là gì?
15. Nêu bốn ranh giới nghề nghiệp không nên vượt dù luật chưa cấm, và giải thích chế tài thực tế của việc vượt qua chúng.
16. Nêu ba quy luật khi đọc một dự báo công nghệ và giải thích nguyên nhân của từng quy luật.

B. Bài tập tính toán

Sáu bài tập dưới đây đều giải được bằng tám công thức của chương, và mỗi bài yêu cầu trình bày phép tính cùng mọi giả thiết chứ không chỉ ghi kết quả.

Bài 1 — thời gian thật cho một cảnh dùng được

Một công cụ tạo sinh mất 45 giây để sinh một cảnh 6 giây, còn người dựng mất 90 giây để xem và quyết định giữ hay bỏ. Qua đo đạc trên mười việc điển hình, công cụ sinh tất cả 40 lần và giữ được 10 cảnh.

- a) Tính tỉ lệ kết quả dùng được η.
- b) Tính thời gian thật cho một cảnh dùng được.
- c) Nếu làm thủ công mất 20 phút một cảnh thì phương án nào nhanh hơn?
- *Đáp số: a) η = 25 %. b) t_{eff} = 540 giây, tức 9 phút 0 giây. c) làm thủ công mất 1.200 giây, nên phương án AI vẫn nhanh hơn, nhưng chỉ nhanh hơn 2,22 lần chứ không phải 27 lần như con số quảng cáo gợi ra.*

Bài 2 — điểm hoà vốn của một khoản đầu tư

Một cơ sở định mua giấy phép một công cụ khử ồn với chi phí thiết lập (gồm giấy phép và học nghề) là 8 triệu đồng. Làm thủ công một đoạn mất 40 phút với đơn giá 150 nghìn đồng một giờ, trong khi công cụ sinh kết quả mất 2 nghìn đồng một lần, người kiểm mất 3 phút và η đo được là 0,7.

- a) Tính chi phí làm thủ công một lần và chi phí thực của một kết quả bằng công cụ.
- b) Tính điểm hoà vốn.
- c) Cơ sở dự kiến làm 150 lần trong năm. Có nên đầu tư hay không?
- *Đáp số: a) làm thủ công 100.000 đồng, công cụ 13.571 đồng. b) N* = 93 lần.* c) 150 lần vượt điểm hoà vốn, nên khoản đầu tư là hợp lý.*

Bài 3 — khi công cụ không rẻ hơn làm tay

Với cùng cơ sở ấy nhưng ở một loại việc khó hơn, η chỉ đạt 0,08 và người kiểm mất 8 phút mỗi lần.

- a) Tính lại chi phí thật một kết quả.

- b) Tính điểm hoà vốn và giải thích kết quả nhận được.
- c) Rút ra kết luận về vai trò của η.
- *Đáp số: a) 275.000 đồng, tức đắt hơn làm thủ công (100.000 đồng). b) không có điểm hoà vốn vì công thức cho kết quả vô cực do mẫu số âm; "không nên dùng" là kết luận hợp lệ và chính xác trong trường hợp này. c) với cùng một công cụ và cùng một giá, chỉ riêng η thay đổi đã làm kết luận đảo chiều hoàn toàn.*

Bài 4 — thời gian sinh và dung lượng kho

Một phim ngắn có 120 cảnh, mỗi cảnh dài 6 giây ở 24 hình một giây. Công cụ sinh được 3 khung hình một giây và mỗi phương án nặng theo tốc độ bit 60 Mbit/s. Kế hoạch là sinh 8 phương án cho mỗi cảnh và giữ lại 25 %.

- a) Tính số khung hình và thời gian sinh một phương án của một cảnh.
- b) Tính tổng thời gian máy phải chạy cho cả phim.
- c) Tính dung lượng kho lúc cao điểm và sau khi dọn.
- *Đáp số: a) 144 khung, 48 giây. b) 46.080 giây, tức 12 giờ 48 phút máy chạy liên tục. c) cao điểm 43,2 GB, sau khi dọn 10,8 GB.*

Bài 5: ngân sách độ trễ cho một buổi dựng từ xa

Một phòng dựng ở Hà Nội làm việc trên tư liệu đặt tại một trung tâm dữ liệu, với các số đo như sau: đường mạng đi và về 60 ms, giải mã ở máy người dùng 30 ms, hiển thị 20 ms.

- a) Tính độ trễ vòng và quy ra số khung hình ở 24 hình một giây.
- b) So với ba ngưỡng cảm nhận, buổi dựng này thuộc mức nào?
- c) Nếu chuyển sang trung tâm dữ liệu ở châu lục khác thì đường mạng thành 200 ms. Kết luận thay đổi như thế nào?
- *Đáp số: a) 110 ms, khoảng 2,6 khung hình. b) trên ngưỡng 100 ms nhưng dưới 200 ms, tức là nhận thấy độ trễ nhưng vẫn làm việc được. c) 250 ms, đã vượt ngưỡng 200 ms và tiến gần ngưỡng khó chịu, nên chỉ phù hợp để duyệt chứ không phù hợp để dựng.*

Bài 6 — lợi ích của một bộ mã hoá mới

Một bộ mã hoá dùng mạng thần kinh đạt chất lượng nghe tương đương ở 4 kbit/s, trong khi bộ mã hoá truyền thống cần 32 kbit/s cho cùng nội dung tiếng nói. Với nhạc, hai con số tương ứng là 96 và 40 kbit/s.

- a) Tính tỉ số nén và phần trăm tiết kiệm cho tiếng nói.
- b) Tính hai con số ấy cho nhạc.
- c) Giải thích vì sao hai kết quả chênh nhau nhiều như vậy.
- *Đáp số: a) r = 8 lần, σ = 87,5 %. b) r = 2,4 lần, σ = 58,3 %. c) tiếng nói là tín hiệu có tính quy luật cao nên mạng học được và chỉ cần truyền phần sai khác, trong khi nhạc có tính quy luật thấp hơn nhiều.*

C. Bài thảo luận

Tình huống 1: tư liệu tài liệu được nâng độ phân giải

Một đài truyền hình làm phim tài liệu kỷ niệm, có sử dụng khoảng mười phút tư liệu lưu trữ từ những năm 1970 vốn mờ và nhiễu. Một biên tập viên đề nghị dùng công cụ nâng độ phân giải để "trả lại chất lượng vốn có" cho tư liệu.

- Cụm từ "trả lại chất lượng vốn có" đúng ở điểm nào và chưa chính xác ở điểm nào?
- Nêu các phương án xử lý, từ không can thiệp tới nâng cấp tối đa, và hệ quả của từng phương án đối với giá trị của tư liệu lịch sử.
- Nếu quyết định nâng cấp thì cần làm thêm những gì để giữ được tính trung thực, cụ thể là những việc phải ghi và những việc phải lưu?
- Nghĩa vụ pháp lý nào phát sinh trong trường hợp này? Áp dụng sơ đồ ba câu hỏi ở mục 12.22.4 và trình bày kết luận.

Tình huống 2: một diễn viên không thu lại được thoại

Một phim đã quay xong, trong đó một diễn viên phụ có ba câu thoại thu hỏng và người ấy đã đi công tác dài ngày nên không thu lại được. Kỹ thuật viên tiếng đề nghị nhân bản giọng từ những câu đã thu tốt trong chính bộ phim để dựng lại ba câu ấy.

- Về mặt kỹ thuật, việc này khả thi đến mức nào và cần những điều kiện gì?
- Về mặt quyền nhân thân, cần những gì trước khi thực hiện, và hợp đồng diễn xuất đã ký cho phim này có đủ hay không? Vì sao?
- Soạn khung một văn bản đồng ý cho trường hợp cụ thể này, đủ sáu điểm ở mục 12.25.2.
- Nếu không liên lạc được với diễn viên trước hạn giao phim thì nên xử lý thế nào? Nêu ít nhất hai phương án thay thế và so sánh chúng.

Tình huống 3 — một cơ sở đào tạo cân nhắc đầu tư

Một cơ sở đào tạo có một khoản kinh phí và ba đề xuất chi: mua giấy phép ba năm cho một bộ công cụ tạo sinh; lắp một trường quay ảo nhỏ; hoặc nâng cấp hệ thống lưu trữ và đường truyền cho phòng dựng.

- Với mỗi đề xuất, áp dụng sáu câu hỏi ở mục 12.6 và nêu kết luận.
- Đề xuất nào chịu rủi ro lạc hậu cao nhất trong ba năm? Lập luận bằng ba bậc tuổi thọ vật liệu ở mục 12.1.
- Nếu chỉ chọn được một đề xuất thì chọn đề xuất nào? Trình bày lập luận bằng khái niệm điểm hoà vốn, với số lần sử dụng ước tính của một cơ sở đào tạo.
- Nêu ba câu hỏi cần đặt ra cho đơn vị cung cấp trước khi ký, và giải thích vì sao ba câu hỏi ấy có ý nghĩa quan trọng hơn cả.

Tình huống 4 — một dự báo mới xuất hiện

Một báo cáo ngành vừa công bố dự báo rằng trong năm năm tới, phần lớn phim quảng cáo sẽ được sản xuất hoàn toàn bằng công cụ tạo sinh và nghề quay phim quảng cáo sẽ thu hẹp mạnh.

- Tách dự báo ấy thành ba loại phát biểu theo mục 12.27.3 và gán cho mỗi loại một mức tin cậy.

- Những điều kiện nào phải xảy ra để phần dự báo về năng lực thành sự thật? Những điều kiện ấy hiện đứng ở đâu?
- Phần dự báo về thời điểm dựa vào giả thiết nào? Giả thiết ấy có tính tới yếu tố nào ngoài kỹ thuật không?
- Đối với một người mới vào nghề quay phim, dự báo này nên dẫn tới việc điều chỉnh những gì và không nên dẫn tới việc điều chỉnh những gì?

TÀI LIỆU THAM KHẢO CHƯƠNG 12

[1] Luật số 134/2025/QH15 — Luật Trí tuệ nhân tạo. Thông qua ngày 10/12/2025, có hiệu lực từ 01/3/2026. Nguồn của toàn bộ mục 12.21 và 12.22; trọng tâm là Điều 11 về minh bạch.

[2] Quy định (EU) 2024/1689 — Đạo luật Trí tuệ nhân tạo của Liên minh châu Âu. Điều 50 về nghĩa vụ minh bạch, áp dụng từ 02/8/2026. Nguồn của mục 12.23.

[3] Luật Sở hữu trí tuệ hiện hành của Việt Nam và các văn bản hướng dẫn thi hành — cơ sở của mục 12.26 về quyền tác giả.

[4] Bộ luật Dân sự hiện hành — các quy định về quyền của cá nhân đối với hình ảnh, là cơ sở trong nước của mục 12.25.

[5] SMPTE ST 2110 — Professional Media Over Managed IP Networks. Nguồn của phần hạ tầng IP ở mục 12.18.

[6] Đặc tả kỹ thuật về xuất xứ nội dung của liên minh C2PA — Coalition for Content Provenance and Authenticity. Nguồn của mục 12.24 về bản kê xuất xứ ký số.

[7] ITU-R BT.2100 và SMPTE ST 2084 — đã dẫn ở Chương 11, được nhắc lại trong phần sản xuất ảo ở mục 12.15.

[8] Các thoả ước lao động tập thể của tổ chức nghề nghiệp diễn viên ở thị trường nước ngoài, phần quy định về bản sao số của người biểu diễn. Nguồn của sáu điểm bắt buộc ở mục 12.25.2. Đây không phải văn bản pháp luật Việt Nam, nhưng là nơi vấn đề đã được thương lượng kỹ nhất cho tới nay.

[9] Nghị định số 142/2026/NĐ-CP ngày 30/4/2026 của Chính phủ hướng dẫn Luật Trí tuệ nhân tạo. Điều 18 quy định thông báo và gắn nhãn nội dung do hệ thống trí tuệ nhân tạo tạo ra hoặc chỉnh sửa. Nguồn của mục 12.21.2.

[10] Gói sửa đổi "Digital Omnibus on AI" của Liên minh châu Âu đối với Quy định (EU) 2024/1689 — tra bản đã công bố trên Công báo của Liên minh. Nguồn của các mốc thời gian ở mục 12.21.2 và 12.23.2.

PHỤ LỤC. DANH MỤC CÔNG CỤ VÀ THIẾT BỊ THAM KHẢO (TÍNH ĐẾN THÁNG 11 NĂM 2026)

Phụ lục này liệt kê các công cụ và thiết bị tiêu biểu của từng nhóm đã trình bày trong chương. Đây là vật liệu có tuổi thọ ngắn nhất của chương, nên phụ lục được tách riêng để cập nhật hằng năm mà không phải sửa phần thân. Danh mục

chỉ nêu tên và vai trò; không nêu giá, vì giá và mô hình thu tiền thay đổi theo quý. Trước khi dùng một công cụ cho dự án thật, người đọc đối chiếu thông tin trên trang chính thức của nhà phát hành và áp dụng sáu câu hỏi đánh giá ở mục 12.6.

Bảng PL.1 — Công cụ tạo sinh video và hình ảnh

Công cụ (tính đến 11/2026)	Nhóm tác vụ	Nhận xét khi đánh giá
Veo (Google)	Tạo sinh video từ văn bản hoặc ảnh	Được nhiều đánh giá nêu là có độ phân giải cao và tiếng đi kèm; cần kiểm tra điều kiện sử dụng thương mại
Seedance (ByteDance)	Tạo sinh video từ văn bản hoặc ảnh	Thường xuất hiện ở nhóm đầu các bảng xếp hạng so sánh công khai; phiên bản đổi nhanh
Kling (Kuaishou)	Tạo sinh video, có công cụ điều khiển chuyển động	Được nhắc đến nhiều trong quy trình cần kiểm soát khuôn hình
Runway	Tạo sinh và biến đổi video, có công cụ cho người dựng	Hướng tới người làm nghề; cần xem tài liệu của chính nhà phát hành
Firefly (Adobe)	Tạo sinh ảnh và video, mở rộng cảnh, xoá vật thể	Tích hợp trong phần mềm của cùng hãng; cần xem điều khoản về dữ liệu huấn luyện

Một số công cụ trong nhóm này đã đổi tên, đổi gói hoặc ngừng cung cấp cho công chúng trong vòng một năm; vì vậy thứ tự trong bảng không phản ánh xếp hạng và không phải khuyến nghị mua.

Bảng PL.2 — Phần mềm dựng và hậu kỳ hình có tính năng AI

Phần mềm (tính đến 11/2026)	Tính năng AI đáng chú ý	Mức can thiệp thường gặp
DaVinci Resolve 21 (ra mắt tại NAB 2026, bản chính thức đầu tháng 6/2026)	Tìm kiếm tư liệu theo nội dung (IntelliSearch), dựng theo kịch bản (IntelliScript), chỉnh tiêu điểm (CineFocus), đọc thông tin bảng ghi cảnh (Slate ID), xoá khuyết điểm trên da	Trợ giúp và bán tự động; kết quả luôn được người dựng duyệt
Premiere (Adobe)	Mở rộng khung hình tạo sinh, tìm kiếm tư liệu, chuyển giọng nói thành phụ đề	Trợ giúp và bán tự động
Các bộ nâng độ phân giải, khử nhiễu hình ảnh	Nâng độ phân giải, khử nhiễu, nội suy khung hình	Biến đổi; áp dụng quy tắc giữ bản gốc chưa xử lý

Bảng PL.3 — Công cụ xử lý tiếng có tính năng AI

Công cụ (tính đến 11/2026)	Nhóm tác vụ	Nhận xét khi đánh giá
iZotope RX 12 (Dialogue Isolate)	Dọn thoại, tách thoại khỏi tiếng nền	Được nhiều bài so sánh xếp ở nhóm giữ được âm sắc gốc của giọng
Adobe Enhance Speech	Làm sạch thoại	Dễ dùng; một số bài so sánh ghi nhận giọng có thể bị đổi màu sắc
LALAL.AI	Tách nguồn (stem) từ bản trộn	Chuyên về tách nguồn, không phải xử lý tổng quát
ElevenLabs	Tạo giọng, nhân bản giọng, tách thoại	Phát sinh nghĩa vụ đồng ý bằng văn bản khi nhân bản giọng người thật

Bảng PL.4 — Thiết bị XR và kính thông minh

Thiết bị (tính đến 11/2026)	Loại	Ghi chú
Apple Vision Pro	Kính đội đầu thực tế hỗn hợp	Vẫn được bán; hướng đi tiếp theo của hãng chưa rõ
Samsung Galaxy XR	Kính đội đầu trên nền tảng Android XR	Đối thủ trực tiếp của kính đội đầu của Apple
Meta Quest	Kính đội đầu thực tế ảo phổ thông	Chiếm phần lớn lượng bán kính đội đầu
Kính thông minh (nhiều hãng, như Ray-Ban Meta, Ray-Ban Display)	Kính có camera, âm thanh hoặc màn hiển thị nhỏ	Thị trường đang dịch từ kính đội đầu sang kính thông minh

Cách dùng phụ lục: mỗi năm nên rà lại từng bảng theo bốn bước. Trước hết kiểm tra mỗi công cụ còn được cung cấp và còn đúng nhóm tác vụ hay không. Tiếp theo đo lại hai số đo của mục 12.5 trên một bộ cảnh chuẩn. Sau đó ghi nhận thay đổi về điều khoản sử dụng, nhất là phần dữ liệu huấn luyện và quyền sử dụng thương mại. Cuối cùng cập nhật cột nhận xét. Việc phát hiện một dòng đã lạc hậu cũng là một bài học: chính nó minh hoạ cho lập luận về vật liệu viết ở đầu chương.