



Integration of AI-powered tools and methodologies in the VET sector

BLOCK B

Generative AI Tools for Text, Audio & Visual Content

Modules 5–8

Work Package 2 - Training Modules Development

Erasmus+ · Pr Nr.: 2025-1-DE02-KA210-VET-000364083



Co-funded by
the European Union



Emphasys
TRAINING CENTRE



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Pr Nr.: 2025-1-DE02-KA210-VET-000364083

Generative AI Tools for Text, Audio & Visual Content

About the Project

AI in ACTION – Integration of AI-powered tools and methodologies in the VET sector is an Erasmus+ project supporting Vocational Education and Training (VET) professionals in understanding and effectively using Artificial Intelligence. It focuses on making AI practical and accessible for teaching, learning, communication, content creation and everyday professional tasks.

The project combines essential AI knowledge with hands-on activities and real VET examples, helping participants use AI tools effectively, understand their possibilities and limitations, critically evaluate their outputs, and integrate AI responsibly into their work.

Who is it for?

The training programme is designed for VET educators, trainers, mentors and administrators, including those with little or no previous AI experience. It progressively develops participants' knowledge, practical skills and confidence through tasks connected to real VET contexts.

The Training Programme

The AI in ACTION training programme consists of 16 modules divided into 4 thematic Blocks, with each module designed as a 90-minute learning session. The programme progresses from AI foundations to content creation, professional applications, and the responsible and sustainable use of AI.

Participants combine theoretical understanding with practical exercises, demonstrations, discussions and hands-on tasks. They learn how to select and apply AI tools for different purposes, evaluate AI-generated outputs and transfer what they learn to their own professional environments.



The four Blocks are:

Block A – Foundations of AI and Large Language Models

Introduces the foundations of AI, Large Language Models and effective prompting.

Block B – Generative AI Tools for Text, Audio & Visual Content

Moves into hands-on production using AI for professional text, voice and narration, music, video and visual content.

Block C – AI Applications in Professional and Personal Projects

Explores local AI models, agentic workflows, AI-powered data analysis, value creation and the development of participants' own AI implementation projects.

Block D – AI Strategy, Ethics, Funding and Sustainability

Addresses responsible AI use, deepfakes and misinformation, environmental sustainability, and the impact of AI on future skills and career pathways.

Together, the four Blocks create a **step-by-step learning journey from understanding AI to applying it in practice and integrating it responsibly into professional work.**

About Block B

Block B focuses on the practical use of Generative AI for creating professional text, audio and visual content. Building on the foundations of Block A, participants explore AI tools for professional writing, voice and narration, music and sound generation, and video and visual content through hands-on activities.

By the end of the Block, participants will have created a **portfolio of AI-generated materials** and developed practical skills for improving their outputs through prompting, experimentation and iterative refinement.

Block B includes four modules:

Module 5: Mastering Text Generation - LLMs and Advanced Prompting

Module 6: Creating Professional AI Voice & Narration

Module 7: AI-Powered Music and Sound Generation

Module 8: Text-to-Video and Visual Content Production



Contents

5

MODULE 5

Mastering Text Generation - LLMs and Advanced Prompting**6**

MODULE 6

Creating Professional AI Voice & Narration**7**

MODULE 7

AI-Powered Music and Sound Generation**8**

MODULE 8

Text-to-Video and Visual Content Production

Generative AI Tools for Text, Audio & Visual Content

Block Overview

Block B is the production block. Building on the foundations and prompting craft of Block A, participants now create real, usable assets: professional text documents, narrated audio in multiple languages, music and soundscapes, and short videos and visual materials they could publish the same week. Every module follows the same rhythm: a compact tool-landscape briefing, a live demonstration, then extended hands-on production time with structured iteration (“first output is rarely the keeper - iterate at least 3 times, save your best prompts”).

Modules	5. Mastering Text Generation - LLMs and Advanced Prompting 6. Creating Professional AI Voice & Narration 7. AI-Powered Music and Sound Generation 8. Text-to-Video and Visual Content Production
Total duration	Each Module is planned for 90 minutes
Target group	VET educators, trainers, mentors, administrators; completion of Block A assumed
Materials needed	Laptop + headphones per participant, microphone (built-in is fine), free accounts prepared in advance: ChatGPT/Claude, ElevenLabs, Suno, Canva, Google AI Studio; one facilitator account on a paid video tool for demo (HeyGen or Runway)
Block outcome	Each participant leaves with a portfolio of self-produced assets: one polished text document, one narrated audio clip (ideally multilingual), one music track, one short video or visual asset - plus the prompts that produced them

Note on tools and pricing: the tool landscape changes fast. All named tools and prices reflect the 2026 landscape at time of writing and should be re-verified before each delivery. The *categories* and *workflows* are stable; specific market leaders rotate.

Consent and ethics rule for the whole block (state at the start of the day): no cloning of anyone’s voice or use of anyone’s face without explicit written consent - including your own colleagues’. Full treatment of risks follows in Block D; this rule applies from the first minute.



BLOCK B · GENERATIVE AI TOOLS FOR TEXT, AUDIO & VISUAL CONTENT

MODULE 5

Mastering Text Generation - LLMs and Advanced Prompting

Mastering Text Generation - LLMs and Advanced Prompting

Module Overview

Duration	90 minutes
Format	Applied writing workshop - minimal input, maximal production
Prerequisites	Module 4 (C.R.A.F.T., few-shot, negative, reverse prompting)
Materials	Laptops, LLM access, participants' own real materials (a lesson plan, a manual excerpt, an email thread - anonymised), Document Production Checklist handout

Learning Objectives

By the end of this module, participants will be able to:

1. Move from single prompts to **multi-turn workflows**: draft → critique → revise → format.
2. Produce four professional text genres with AI support: instructional materials, assessments, workplace communication, and project/funding documents.
3. Use iterative refinement systematically (“never accept the first draft”).
4. Apply style transfer and few-shot techniques to keep AI text in their own voice.
5. Build differentiated versions of one source text for different learner levels and languages.
6. Verify and edit AI text against a quality checklist before it leaves their machine.

Session Plan

Time	Segment	Method
0–10 min	Recap quiz on Block A prompting + warm-up prompt	Plenary
10–25 min	Input: From prompts to workflows - the multi-turn method	Presentation + live demo
25–40 min	Input: The four text genres of VET work, with prompt patterns	Presentation
40–75 min	Practical Task: The Text Production Studio (choose 2 of 4 stations)	Individual/pairs
75–85 min	Peer review with the Document Production Checklist	Pairs
85–90 min	Save to prompt library + wrap-up	Individual

Theoretical Input Part 1: From Prompts to Workflows (15 min)

One good prompt produces a draft. A **workflow** produces a finished document. The professional pattern has four turns:

1. **Draft** - a full C.R.A.F.T. prompt produces version 1.



Mastering Text Generation - LLMs and Advanced Prompting

2. **Critique** - instead of editing yourself immediately, ask the model to attack its own work: “Act as a strict quality reviewer. List the 5 biggest weaknesses of this text for [audience]. Do not rewrite yet.”
3. **Revise** - “Fix weaknesses 1, 3 and 4. Keep everything else unchanged.” (Selective revision keeps control with you.)
4. **Format** - “Now format this as [handout / table / step-by-step checklist] ready to paste into a document.”

Live demonstration (facilitator): run all four turns on a real example - e.g., a workplace safety briefing for a workshop induction. Show how the critique turn surfaces problems (missing PPE items, no accident procedure, reading level too high) that a single prompt would have shipped.

Two supporting habits:

- **Style transfer / few-shot:** paste a text you have written and liked: “Analyse my tone and sentence rhythm. All following outputs must match it.” AI text that sounds like *you* needs less editing and is more honest to reuse.
- **Chunking long documents:** models work best on one section at a time. Outline first (“give me a section outline only”), approve the outline, then generate section by section.

Theoretical Input Part 2: The Four Text Genres of VET Work (15 min)

■ Genre 1 - Instructional materials

Lesson plans, worksheets, step-by-step job instructions, theory summaries. **Key prompt elements:** exact learner profile (age, year, prior knowledge), your curriculum constraints, equipment actually available, and negative constraints (“do not include equipment not listed”). **Signature pattern - differentiation:** “Here is my worksheet on [topic]. Produce three versions: (a) support version with scaffolded steps and glossary, (b) standard, (c) extension version with a transfer problem. Same learning objective in all three.”

■ Genre 2 - Assessment materials

Quiz banks, rubrics, practical exam scenarios, model answers. **Key prompt elements:** question mix and counts, difficulty distribution, answer key placement, observable criteria. **Signature pattern - rubric:** “Create an assessment rubric for [practical task] with 4 levels. Every descriptor must be observable behaviour - no vague words like ‘good’ or ‘adequate’. Format as a table. Include a fail-safe row for safety violations that overrides all other scores.”

■ Genre 3 - Workplace communication

Emails to parents, employers, authorities; announcements; meeting summaries. **Key prompt elements:** relationship and tone calibration (“strict but supportive”, “formal but warm”), what must and must not be said, one style example (few-shot). **Signature pattern - difficult email:** “Draft an email to a training company whose apprentice has missed 3 sessions. Goals: preserve the partnership, state facts neutrally, propose a concrete meeting. Do not apologise on the apprentice’s behalf. Do not threaten. 150 words maximum.”



Mastering Text Generation - LLMs and Advanced Prompting

■ Genre 4 - Project and funding documents

Erasmus+ applications, project reports, dissemination texts. **Key prompt elements:** the actual call/criteria text pasted as context, word limits, evaluator perspective. **Signature pattern - evaluator simulation:** “Here is the evaluation rubric and my draft answer. Act as a strict evaluator: score my draft against each criterion and justify each score before suggesting improvements.” **Tip:** for polished EU-register prose in EU languages, run the final pass through DeepL Write.

Practical Task: The Text Production Studio (35 min)

Four stations; each participant completes **two**, using their own real (anonymised) material where possible. Every station requires the full 4-turn workflow - draft, critique, revise, format - and at least 3 iterations total.

Station A - The Differentiated Worksheet. Take one existing worksheet or theory topic from your subject. Produce the three-level differentiated set (support / standard / extension). Bonus: generate a translated support version in a learner’s home language and check it with DeepL.

Station B - The Complete Assessment Pack. For one practical task you actually assess, produce: a 10-question mixed quiz with answer key, an observable-behaviour rubric, and one scenario-based question (“what would you do if...”). Apply negative prompting to keep it inside your real equipment list.

Station C - The Communication Set. Produce three linked texts around one real situation: an email to the trainee (strict but supportive), an email to the employer/parents (formal, factual), and a short internal note for the file. All three must be consistent in facts but different in tone and length.

Station D - The Project Paragraph. Take one real section from a project application or report you are working on (or use the provided sample call text). Run the evaluator-simulation pattern, then revise until your simulated score improves by at least one level.

Peer Review: Document Production Checklist (10 min)

Swap your best output with a partner. Review against the checklist (also a take-home handout):

- **Accurate?** Facts, names, norms, numbers checked by a human. No invented standards.
- **Right level?** Reading level and terminology match the stated audience.
- **Complete?** Nothing critical missing (safety steps, dates, next actions).
- **Your voice?** It doesn’t read like a machine wrote it for a stranger.
- **Legal/privacy?** No personal data was pasted into a consumer tool; nothing confidential included.
- **Format ready?** Could be pasted into your institutional template right now.

Reviewer marks each item ✓/✗ and writes one improvement suggestion. Author gets 3 minutes to apply it.



BLOCK B · GENERATIVE AI TOOLS FOR TEXT, AUDIO & VISUAL CONTENT

MODULE 6

Creating Professional AI Voice & Narration

Creating Professional AI Voice & Narration

Module Overview

Duration	90 minutes
Format	Tool landscape briefing + guided audio production workshop
Prerequisites	Modules 4–5; consent rule from block introduction
Materials	Laptops + headphones, microphones, free accounts: ElevenLabs, Google AI Studio; sample scripts printed; Consent Form template handout

Learning Objectives

By the end of this module, participants will be able to:

1. Map the 2026 sound/voice tool landscape across five categories: voice generation, voice cloning, transcription, translation + dubbing, and integrated multimodal workflows.
2. Produce professional narration from text in multiple voices and languages.
3. Clone their own voice (with consent protocol) and evaluate quality and risks.
4. Transcribe and translate spoken content and re-voice it in another language.
5. Identify concrete accessibility and inclusion applications for VET learners (reading difficulties, multilingual groups, distance learners).
6. Apply the consent and disclosure rules that govern any use of synthetic voice.

Session Plan

Time	Segment	Method
0–10 min	Listening test: real or AI voice? + discussion	Plenary
10–25 min	Input: The Sound-Gen landscape 2026	Presentation + live demos
25–35 min	Input: Voice cloning - how, why, and the consent protocol	Presentation
35–75 min	Practical Task: The Audio Studio (choose 1 main project)	Individual/pairs
75–85 min	Debrief: what surprised you, what scared you?	Plenary
85–90 min	Accessibility applications + wrap-up	Plenary

Opening: The Listening Test (10 min)

Play 4 short clips prepared in advance: two human recordings, two AI-generated (one obviously synthetic older TTS, one modern ElevenLabs-quality). Participants vote real/AI for each. Reveal answers. **Expected result:** the modern clip fools most of the room. **Discussion:** if you can't hear the difference, what follows - for accessibility (good) and for fraud and deepfakes (bad, treated fully in Module 14)?



Creating Professional AI Voice & Narration

Theoretical Input Part 1: The Sound-Gen Landscape 2026 (15 min)

Category	What it does	Leading tools (2026)	Typical paid tier
Voice generation	Realistic narration in any voice + language from text	ElevenLabs v3 · Google AI Studio (Gemini native audio) · Cartesia	ElevenLabs Starter ~\$5/mo; Cartesia Pro from ~\$5/mo
Voice cloning	Your own voice speaking any text	ElevenLabs Voice	ElevenLabs Creator ~\$22/mo (requires consent verification)
Transcription	Audio/video → searchable text	Whisper v3 (free, open-source) · ElevenLabs · Trint (strong, no free tier)	Otter Pro ~\$17/mo; free path: Whisper via Hugging Face
Translation + dubbing	Re-voice content in another language	ElevenLabs · YouTube auto-dub · CapCut · HeyGen	ElevenLabs Creator; HeyGen Creator ~\$24/mo
Integrated multimodal	One workflow: upload audio → transcribe → translate → re-voice	Google AI Studio (Gemini 2.5 Native Audio) - free	Free at time of writing

Live demonstrations (pick two): (a) type one sentence into ElevenLabs, generate it in three voices and two languages; (b) upload a 20-second recording to Google AI Studio and run transcribe → translate → re-voice in one session.

Practical guidance: free tiers cover exploration and most classroom use; pay only when a clear weekly task pays back the fee (the recurring rule of Block B).

Theoretical Input Part 2: Voice Cloning and the Consent Protocol (10 min)

How it works: modern systems need only 1–3 minutes of clean recorded speech to build a usable clone of a voice - timbre, accent and rhythm included.

Legitimate VET uses: narrating your own course materials at scale without recording every update; keeping one consistent “course voice” across modules; multilingual narration in your own voice for international learners; accessibility narration for trainees with reading difficulties.

The Consent Protocol (non-negotiable, handout provided):

- 1. Written consent** from the voice owner before any cloning - including yourself as employee (check institutional policy).
- 2. Stated purpose and scope:** what content, which platforms, for how long.
- 3. Revocability:** the voice owner can withdraw consent; clones are then deleted.
- 4. Disclosure:** synthetic narration is labelled as such wherever practical.
- 5. Storage:** clone access credentials treated like personal data under GDPR.

Discussion prompt: several EU jurisdictions are tightening voice-likeness law. Who in your institution would need to approve a cloned-voice course narration? (Most rooms cannot answer - that’s the point; institutional policy is a Block D theme.)



Creating Professional AI Voice & Narration

Practical Task: The Audio Studio (40 min)

Pick **one** main project. Iterate at least 3 times. Save your best prompt/settings and the final audio file. Work in pairs if accounts are limited.

Project A - Accessibility narration (recommended default). Tool: ElevenLabs v3 (+ your voice, with signed consent form). Clone your own voice. Narrate a 30-second excerpt from your actual course material in your voice - then generate the same excerpt in two other languages. Test: imagine a learner with reading difficulty using this. What would you produce first for them?

Project B - Podcast intro package. Tool: ElevenLabs v3 (+ Suno if time allows - full music treatment in Module 7). Script and voice a 15-second intro for a recurring “Subject of the Week” classroom podcast: “Welcome to [your subject] - let’s get into it.” Produce two voice options (one energetic, one calm) and decide which fits your learners.

Project C - Multilingual briefing. Tool: Google AI Studio (Gemini 2.5 Native Audio) - free. Record yourself reading a real 30-second announcement (open day, exam dates). Transcribe → translate into Greek or German (or a learner language in your institution) → re-voice. Compare naturalness across languages; note where technical terms break.

Project D - Pronunciation pack for language/integration classes. Tool: ElevenLabs. Generate a set of 10 short sentences containing your subject’s core technical vocabulary, spoken slowly and clearly, then at natural speed. This becomes a self-study resource for non-native-speaker trainees.

Further ideas for fast finishers: school radio announcement · guided 2-minute relaxation for exam stress · audiobook narration of a textbook section · translated welcome message for international students · voice-over for an existing tutorial video.

Debrief: What Surprised You, What Scared You? (10 min)

Structured plenary round:

- How realistic was the cloned/generated voice? Where did it fail (emotion, technical terms, your language)?
- Could this assist your training concretely - name the first real use.
- What is the risk twin of your use case? (Every legitimate use has an abuse mirror - collect these on flipchart; they seed Module 14.)

Closing: Accessibility as the Killer Application (5 min)

Frame the module’s takeaway: for VET, the strongest case for AI voice is **inclusion** - learners with dyslexia or reading difficulties, learners whose first language differs from the language of instruction, and distance/shift learners who consume content while commuting or working. A narrated version of every key handout used to be economically impossible. It now costs minutes.



BLOCK B · GENERATIVE AI TOOLS FOR TEXT, AUDIO & VISUAL CONTENT

MODULE 7

AI-Powered Music and Sound Generation

AI-Powered Music and Sound Generation

Module Overview

Duration	90 minutes
Format	Short landscape input + extended creative production + licensing clinic
Prerequisites	Module 6 (accounts and audio handling already set up)
Materials	Laptops + headphones, free accounts: Suno (and/or Udio), one shared example slideshow needing music, Licensing Quick-Guide handout

Learning Objectives

By the end of this module, participants will be able to:

1. Explain how text-to-music generation works at user level (prompt → style, mood, structure, lyrics).
2. Produce royalty-conscious music for concrete educational scenarios: event background, podcast themes, video soundtracks, classroom ambience.
3. Write effective music prompts controlling genre, mood, tempo, instrumentation and structure.
4. Iterate on generated music: extending, remixing mood variants, and matching music length to content.
5. Read a tool's licensing terms and decide whether an output may be used commercially / publicly / institutionally.
6. Identify where music genuinely improves learning materials - and where it distracts.

Session Plan

Time	Segment	Method
0–10 min	Demo: one prompt → one song, live	Plenary
10–25 min	Input: How music generation works + prompt anatomy	Presentation
25–35 min	Input: Licensing - who owns AI music?	Presentation + real ToS reading
35–70 min	Practical Task: The Sound Studio (choose 1 main project)	Individual/pairs
70–85 min	Listening session + peer vote	Plenary
85–90 min	"Does it help learning?" reflection + wrap-up	Plenary

Opening Demo (10 min)

Live, in front of the room: prompt Suno with “uplifting corporate-folk instrumental, acoustic guitar and light percussion, 60 seconds, optimistic build, no vocals - background music for a vocational school open-day slideshow.” Play the result. Then change one word (“optimistic” → “nostalgic”) and regenerate. The mood shift in one word is the core lesson: **you direct; the model performs.**



AI-Powered Music and Sound Generation

Theoretical Input Part 1: Prompt Anatomy for Music (15 min)

Leading tools in 2026: **Suno v4.5** (most popular, strong vocals and structure; Pro ~\$10/mo adds commercial licence, faster generation, longer tracks) and **Udio** (Standard ~\$10/mo, comparable quality). Both work from a text prompt plus optional lyrics.

The five controls of a music prompt:

1. **Genre/style:** “acoustic folk”, “synthwave”, “brass band”, “lo-fi hip-hop”. Combining two genres often produces the most usable, non-generic results.
2. **Mood:** “hopeful and modern”, “warm and nostalgic”, “focused and calm”. The single highest-impact word in the prompt.
3. **Tempo/energy:** “slow build”, “steady 90 bpm”, “energetic chorus”.
4. **Instrumentation:** name instruments to include - and use negative prompting to exclude (“no vocals”, “no heavy drums”).
5. **Structure/length:** “15-second jingle”, “60-second loopable background”, “intro–verse–chorus”.

With or without lyrics: for educational use, instrumental is usually right (background music must not compete with speech). Lyrics shine in one niche: **memory aids** - a chorus that encodes a formula, a safety rule or vocabulary is a legitimate mnemonic technique (“sing the resistor colour code”).

Theoretical Input Part 2: Licensing - Who Owns AI Music? (10 min)

Three questions to answer before any AI music leaves your machine:

1. **What does the tool’s licence say?** Typically: free tier = non-commercial personal use only; paid tier = commercial licence to *you*. Educational public use (open days, published videos) usually requires the paid tier. **Exercise:** open Suno’s terms live and find the answer together - reading real ToS is the skill.
2. **Who owns the output?** Contested legal territory: in many jurisdictions purely AI-generated works have weak or no copyright protection, but the *tool’s contract* still binds you. Institutional rule of thumb: treat the licence terms as governing.
3. **Could it infringe?** Models trained on existing music can produce output resembling real artists. Never prompt “in the style of [named artist]” for anything public; describe the style in generic terms instead.

Handout: one-page Licensing Quick-Guide with these three questions and the current Suno/Udio answers, plus a “when in doubt” escalation line to institutional policy.



AI-Powered Music and Sound Generation

Practical Task: The Sound Studio (35 min)

Pick **one** main project. Iterate at least 3 times (change one control per iteration and note the effect). Save your best prompt + final track.

Project A - Event music: open-day slideshow. Generate 60 seconds of calm instrumental background for an open-day slideshow. Then produce one alternative mood version (“hopeful and modern” → “warm and nostalgic”). Play both against the shared sample slideshow and pick the better fit - justify the choice.

Project B - Podcast/channel theme. Compose a 15-second theme tune for a recurring “Subject of the Week” classroom podcast, then add the voiced intro produced in Module 6 on top (any simple editor, e.g., CapCut or Audacity). Deliverable: one combined intro file.

Project C - The Mnemonic Song. Choose one hard-to-memorise item from your subject (a formula, a safety sequence, colour codes, irregular verbs). Write lyrics with your LLM (Module 5 workflow!), then generate a short, catchy song in a genre your trainees would tolerate. Test on a neighbour: can they recall the content after two plays?

Project D - Classroom ambience set. Produce three 60-second loopable ambience tracks for distinct classroom phases: focused individual work, group work energy, and end-of-day wind-down. Name the observable behaviour each is meant to support.

Listening Session (15 min)

Each participant/pair plays their best 30 seconds. The room votes on two axes by show of hands: “would use tomorrow” and “most creative”. Facilitator draws out *why* winners work - almost always: specific mood words, deliberate instrumentation choices, and iteration rather than first-take acceptance.

Closing Reflection: Does It Help Learning? (5 min)

Honest framing to close: music is the most fun and the least pedagogically essential tool of the block. Legitimate wins: event atmosphere, video professionalism, mnemonics, classroom-phase signalling, learner projects (trainees producing their own). Traps: background music over spoken instruction reduces comprehension; novelty wears off; production time can exceed learning value. **Rule: music serves the learning goal, or it gets cut.**



BLOCK B · GENERATIVE AI TOOLS FOR TEXT, AUDIO & VISUAL CONTENT

MODULE 8

Text-to-Video and Visual Content Production

Text-to-Video and Visual Content Production

Module Overview

Duration	90 minutes
Format	Two-part production workshop: visual assets (images, design, infographics) + video
Prerequisites	Modules 5–7 (scripts from Module 5, audio from Modules 6–7 feed directly into video)
Materials	Laptops, free accounts: Canva, Microsoft Designer or Adobe Express, Napkin/Whimsical, Gamma; facilitator demo account on HeyGen or Runway; participants' phones for recording raw footage

Learning Objectives

By the end of this module, participants will be able to:

1. Map the 2026 visual-generation landscape across six categories: image, logo, poster/flyer, infographic, presentation, and visual storytelling.
2. Map the 2026 video-generation landscape across five categories: text-to-video, image-to-video, avatar videos, AI editing, and social/explainer clips.
3. Produce one publication-ready visual asset for a real work purpose (recruitment, dissemination, explanation, celebration).
4. Produce or storyboard one 30-second video using an appropriate tool chain.
5. Write effective visual prompts (subject, style, composition, format/aspect ratio, constraints).
6. Chain tools into a complete production pipeline: script → voice → visuals → edit.
7. Judge output quality against a “good enough to publish?” standard, including brand and bias checks.

Session Plan

Time	Segment	Method
0–15 min	Input: Visual-Gen landscape + visual prompt anatomy	Presentation + demo
15–40 min	Practical Task 1: Make a visual asset for your actual work	Individual/pairs
40–55 min	Input: Video-Gen landscape + the one-afternoon recipe	Presentation + demo
55–80 min	Practical Task 2: The 30-second video	Individual/pairs
80–90 min	Gallery walk + quality debrief	Plenary



Text-to-Video and Visual Content Production

Theoretical Input Part 1: The Visual-Gen Landscape 2026 (15 min)

Most tools are free to try; paid tiers unlock resolution, commercial licence, brand kits and batch generation.

Category	Use case	Leading tools (2026)	Paid pick
Image	Photoreal, illustration, concept art	Ideogram 3 · GPT-Image (ChatGPT) · Flux (Black Forest) · Google Imagen 4	Midjourney ~\$10/mo · Ideogram ~\$7/mo
Logo	Brand marks, monograms, icons	Ideogram 3 (best text-in-image) · Recraft · Canva AI (simplest) · Adobe (free tier)	Recraft Pro ~\$12/mo · Looka one-time ~\$20
Poster/Flyer	Workshop posters, event flyers	Canva Magic Studio · Microsoft Designer · Adobe Express · Recraft designer mode	Canva Pro ~€11/mo · Adobe Express ~€10/mo
Infographic	Process flows, comparisons, data viz	Napkin AI · Whimsical AI · Venngage AI · Mermaid Chart · Excalidraw AI	Whimsical ~\$10/mo · Napkin Pro
Presentation	Decks from a prompt or outline	Gamma 2.0 · Decktopus · Beautiful.ai · Plus AI for Google Slides · Claude	Gamma Plus ~\$10/mo · Decktopus Pro ~\$14/mo
Storytelling	Consistent characters, scenes, panels	Midjourney · Adobe Firefly 3 · GPT-Image	Adobe Firefly · Krea Pro ~\$10/mo

Visual prompt anatomy (the five parts): subject (“a young electrician testing a circuit board”) · style (“clean documentary photography”, “flat illustration”, “watercolour”) · composition and lighting (“close-up, natural window light, shallow depth of field”) · format (“1080×1080 square”, “A4 portrait”) · constraints (“no text in image”, “diverse group”, “no brand logos”). **Bias check habit:** generate people twice and inspect - image models default to stereotypes (gender, ethnicity, age) unless you specify.

Practical Task 1: Make a Visual Asset for Your Actual Work (25 min)

Pick **one**. Iterate at least 3 times. Save your best prompt + final output.

A - Recruitment (Instagram-ready). Tool: Canva Magic Studio or Microsoft Designer. “A bold square graphic recruiting students to my programme. Hero image of [your subject in action], one benefit headline, warm but professional tone, 1080×1080.”

B - Dissemination (parents’ open day deck). Tool: Gamma 2.0 or Decktopus. “Generate a 6-slide deck introducing [our programme / the AI in ACTION project] to newcomers: what it is, why it matters, what students and educators gain, one slide on responsible AI.”

C - Visual explainer (infographic). Tool: Napkin AI or Whimsical AI. “Turn [a process from your subject] into a clean infographic - one diagram, one rule, two examples, one common mistake. Educational, approachable, A4 portrait.”

D - Celebration (certificate). Tool: Adobe Express or Canva. “Design an end-of-course certificate template - soft watercolour background, elegant serif typography, space for trainee name, module title, date and signature.” Apply your institution’s branding.



Text-to-Video and Visual Content Production

Theoretical Input Part 2: The Video-Gen Landscape 2026 (15 min)

Category	Use case	Leading tools (2026)	Paid pick
Text-to-video	Prompt → moving footage (often with audio)	Google Veo 3 · Runway Gen-4.5 · Adobe Firefly video	Runway Standard ~\$12/mo
Image-to-video	Bring a still photo to life	Kling 2.1 · Runway	Kling Standard ~\$7/mo (best quality/price) · Kling Pro ~\$25/mo
Avatar videos	A talking head reads your script in any language	HeyGen · Synthesia 2.0	HeyGen Creator ~\$24/mo · Synthesia Starter ~€27/mo
AI editing	Cut, caption, repurpose existing footage	CapCut AI · Descript	Descript Creator ~\$16/mo ("edit video like text")
Explainer/social	End-to-end short promo videos	InVideo AI · Pictory · Synthesia	Synthesia Starter ~€27/mo

The one-afternoon recipe (show as a pipeline diagram): Script (ChatGPT/Claude - Module 5) → Avatar or narration (HeyGen / ElevenLabs - Module 6) → B-roll (Veo 3 / Kling) → Music (Suno - Module 7) → Edit + captions (CapCut). **A complete promotional video in ~90 minutes, end to end.** This pipeline is the synthesis of the whole block.

Facilitator demo: generate one 8-second clip live (Veo/Runway/Kling - whichever account is available) and one 20-second avatar clip from a two-sentence script.

Practical Task 2: The 30-Second Video (25 min)

Pick **one**. Free tiers constrain some options - storyboarding + partial production is an acceptable deliverable where paid tools would be needed; the pipeline plan is the learning objective.

A - Recruitment (TikTok-style). Tool: InVideo AI or Pictory. "30-second vertical recruitment video for my programme. Quick cuts, upbeat track, captions in 2 languages, end card with logo and 'apply by [date]'. Energy: sports-brand promo, not school brochure."

B - Avatar welcome (multilingual). Tool: HeyGen or Synthesia (facilitator account/demo if needed). Script a 30-second welcome to new students explaining their first week. Generate in English; auto-dub to Greek and German. Test all three on a colleague - where does dubbing break?

C - Cinematic shot (Instagram hero clip). Tool: Veo 3, Runway or Kling. "8-second cinematic shot - slow dolly through a [your workshop/classroom/lab] at golden hour, warm lighting, shallow depth of field, no text. Square 1:1."

D - Equipment explainer (repurpose real footage). Tool: CapCut AI or Descript. Record a 60-second phone walkthrough of any piece of equipment. Let AI auto-caption, cut the "umms", and produce a 30-second vertical version for social media. (This option is fully free and uses real footage - often the most institutionally useful.)

Fast-finisher ideas: animated logo intro · "day in the life" student reel · behind-the-scenes clip of a current project · split-screen explainer of a complex process · end-of-year highlights reel.



Text-to-Video and Visual Content Production

Gallery Walk and Quality Debrief (10 min)

All outputs on screens; participants circulate. Then structured debrief:

- **Quality:** was the output “good enough to share”? What separates the shippable from the almost?
- **Time:** how long did it take versus commissioning a designer/videographer? What is the honest quality trade?
- **Money:** which single paid tool would earn back its subscription in your real workflow - and which is a toy?
- **Ethics preview:** whose faces appeared in your generated visuals? Any real people, any stereotypes? (Bridge to Block D.)

Block B Synthesis: Five Sentences to Take Home

1. **You can make it.** You produced text, voice, music, visual and video assets you could ship Monday. The barrier is gone.
2. **Quality scales with iteration.** First output is rarely the keeper. Iterate 3+ times. Save your best prompts.
3. **Free tools cover most needs.** Pay for ElevenLabs / HeyGen / Canva Pro / Midjourney only when a clear weekly task pays back the fee.
4. **Risks are now operational, not theoretical.** Copyright, likeness, bias, privacy - the Responsible AI Checklist (Block D) applies before anything is published.
5. **The pipeline is the skill.** Individual tools rotate yearly; the workflow - script → voice → visuals → edit, with verification at each step - is durable.





Thank you.

Block B · Generative AI Tools for Text, Audio & Visual Content

KA210-VET · Erasmus+ · Pr Nr.: 2025-1-DE02-KA210-VET-000364083



Co-funded by
the European Union



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.