



## **Efficient and sustainable generation of training videos**

Matthias Rösslein – ScientificIT Empa

25.08.2026

# Challenges with any training



- New master's & graduate students start research every month (~50% annual turnover)
- Most topics should start at **day one** — but they are **not part of a typical curriculum**
- Information and skills rely on ad-hoc sessions or supervisor guidance
- 1-hour intro = basic concepts only, no hands-on practice
- Monthly in-depth training is **unsustainable** → need a scalable solution



# First approach to a solution – Video recording



- **Pros:**

- Quick, but can become expensive
- Personal touch
- Flexible recording schedule

- **Challenges:**

- Time-consuming editing
- Unclear structure or arguments
- Speaking too fast or too slow
- Flat or unengaging voice
- Lack of professional presentation skills
- Poor audio or lighting quality
- **Difficult to maintain/update**



# What we need – a maintainable video



- Script as single source of truth — not the recording
- Edit one sentence, regenerate — no re-shoot
- Same voice, look and quality every time
- Independent of one person's availability
- Modular chapters, replaceable one by one
- Easy handover to colleagues
- Consistent quality across the whole training series



# Steps toward a sustainable solution



## 1. Start with the Script

- Plan the entire presentation in text.
- Organize arguments and structure before any video work.

## 2. Create AI Voice Narration

- Use realistic text-to-speech tools.
- Control pacing, tone, emphasis, and clarity.

## 3. Generate AI Presenter Video

- Upload script & audio to an AI video creation platform.
- Choose professional avatars with natural gestures and perfect lip-sync.

## 4. Assemble & Publish

- Combine voice, visuals, and slides/images into a finished video.
- No manual editing required.
- Fast future updates = edit the script → regenerate → publish.



# 1. Start with a script



- **Start with high-**

- A well-formula
- Target the slide

- **Combine both i**

- Specify tone (e
- audience, and

- **Generate and re**

- Review for acc
- RDM standard
- iterate with tar

- **Finalize and pub**

- Approve the re
- Copy the text into your Markdown editor for versioning and reuse

## 1. Slide

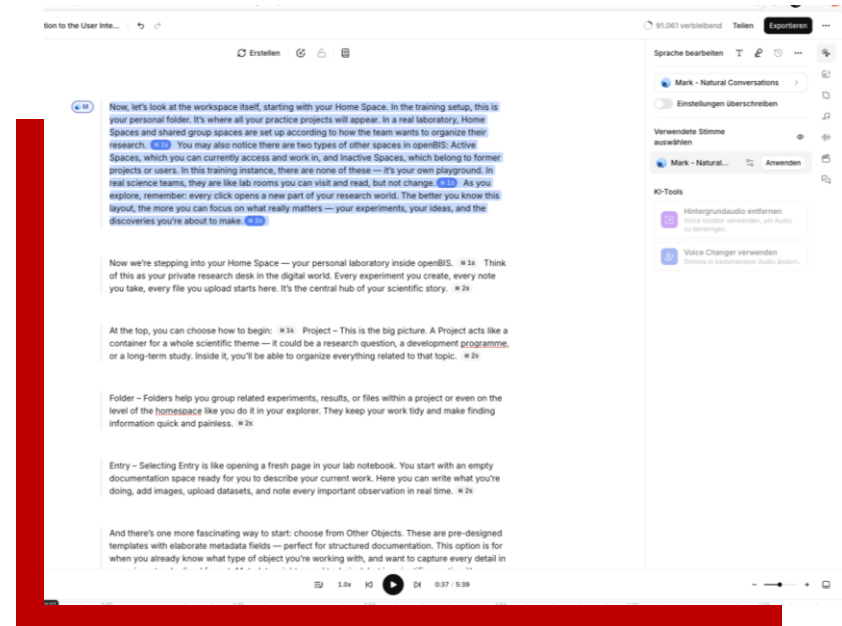
Ready for your first steps in openBIS – the place where your research data finds structure, meaning, and a safe home. / In this session, we'll explore both sides of openBIS. First, the **Eln – the Electronic Laboratory Notebook** – where you record your experiments, keep organised notes, and attach data directly to the work you do. Then, the **Lims – the Laboratory Information Management System** – where samples, materials, and workflows are tracked so every piece of information stays connected and easy to find. / Think of the ELN as your research diary, and the LIMS as your laboratory's memory. Together, they make sure nothing gets lost, and that you can always trace your work from data back to the original experiment. / This is not just a lecture – it's active training. I'll show you each step, and right after, you'll try it yourself on a dedicated training instance. You'll get your hands on the interface, create records, upload data, and see how everything links together in real time. / Why does this matter? Because well-managed data is the backbone of modern materials research. It saves you hours of searching, prevents mistakes, and makes collaboration effortless. And when your results are well-



## 2. AI-Voice & Text to Speech – Elevenlabs



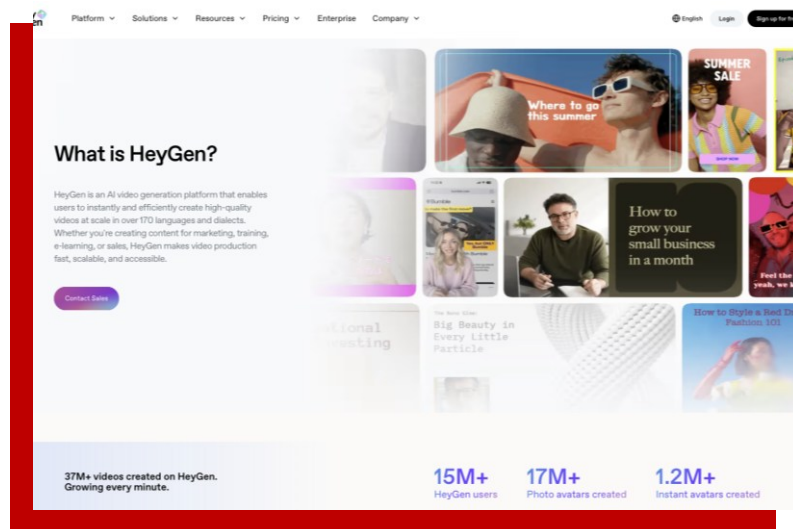
- >10'000+ professional voices in 74 languages for training, teaching, narration, storytelling, etc.
- Each voice model trained on  $\geq 2$  hours of studio-quality recordings
- Supports young, mature, and many different character voices of flexible style
- Paste text → generate high-fidelity MP3s for each slide
- Delivers consistent, natural, and polished audio for any presentation
- Scales easily for projects of any size – from a single lecture to comprehensive e-learning courses



# 3. AI-Powered Video Generation Platform - HeyGen

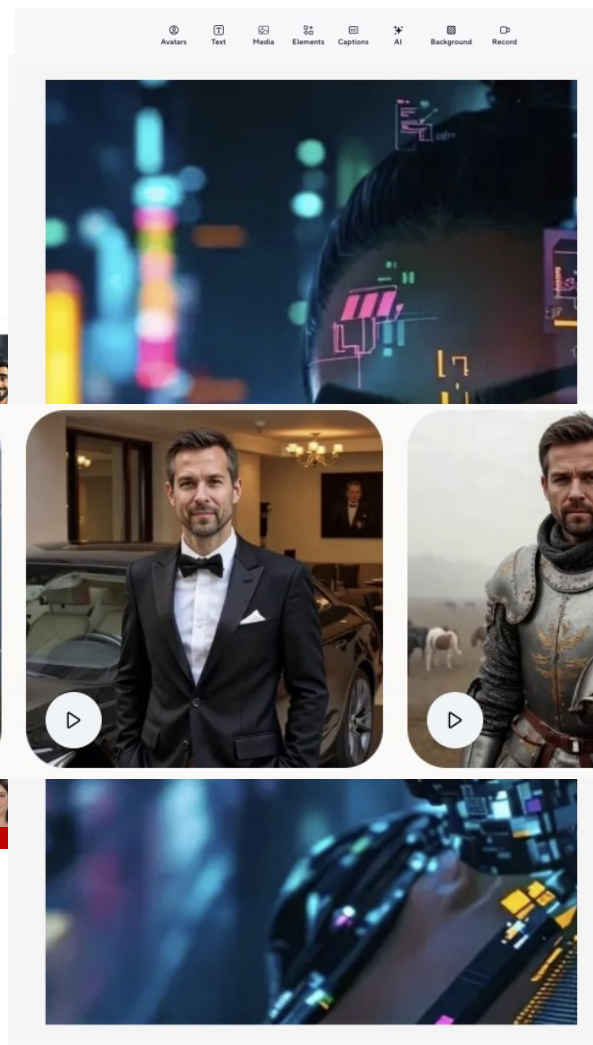
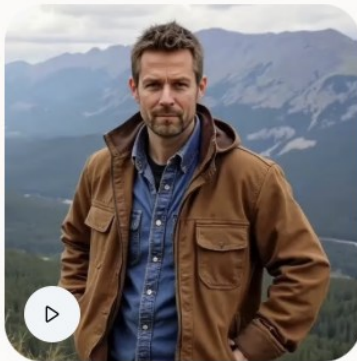
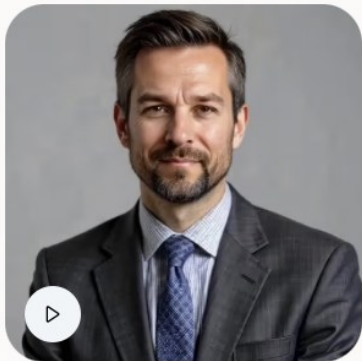


- Founded in 2020 (Joshua Xu & Wayne Liang) to make visual storytelling accessible to all
- Enables fast, low-cost video creation vs. traditional \$1k/minute production costs
- 84'149'070 avatars generated – customizable pitch, speed, and expressions
- New Avatar IV&V feature: create lifelike videos from a single photo or short video sequence in minutes
- Supports marketing, sales, training, product demos, and online education, etc.
- Interactive avatars for real-time conversations, customer support, and virtual meetings
- Serves clients from SMBs to Fortune 500, NGOs, and content creators worldwide
- Positioned in a rapidly growing AI video market projected to reach \$1.9B by 2030
- Competing products are Colossyan, Synthesia and more



# HeyGen – AI Avatar Capabilities

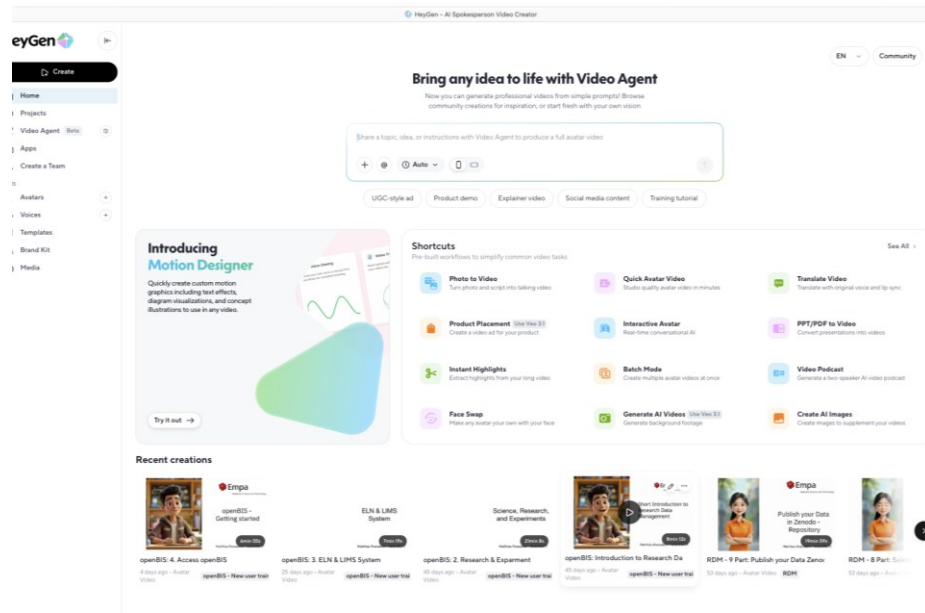
- **Create lifelike avatars in minutes** from text, scripts, or images – no camera, actors, or studio needed
- **Clone yourself** tailored to various audiences
- **Replicate gestural voice tone** for
- **Emotion-adapted** delivery in real time to match your message
- **Fully customizable** – modify pitch, speed, gestures, and appearance to fit brand identity



# 3. HeyGen – AI-Powered Video Generation



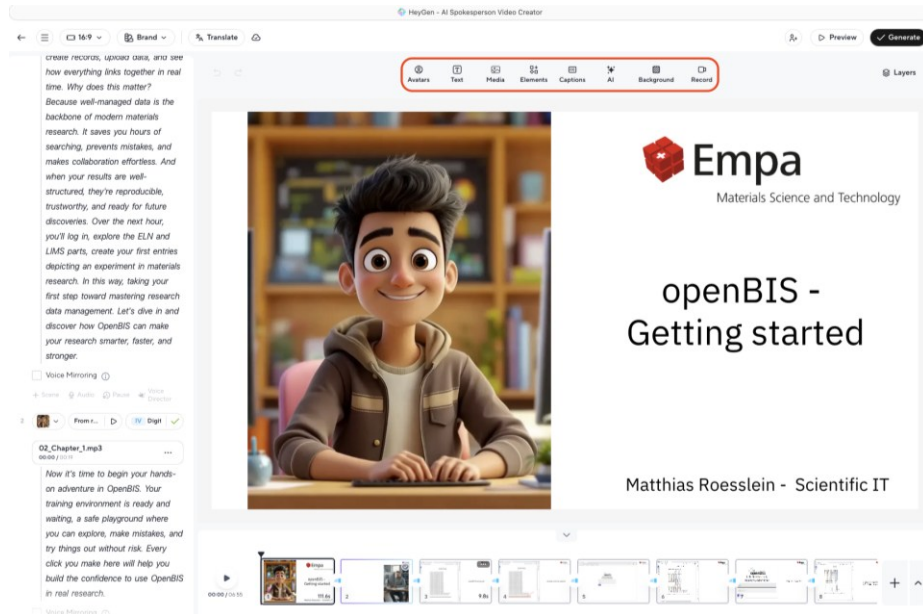
- **Multiple video generation options** – create from avatars, scripts, images, or existing video footage
- **Upload any pictures or videos** to integrate directly into your presentation content
- **Access to leading AI image & video generation tools** for unlimited creative possibilities
- **Drag-and-drop interface** with full editing capabilities for intuitive, fast production
- **Code-based output ensures reproducibility**, making it easy to recreate the same video later
- **Quickly modify or replace individual segments** without re-producing the entire video
- **Ideal for updates, version control, and iterative improvements** in marketing, training, or creative work



# 3. HeyGen – PDF/PowerPoint to AI Video Workflow



- **Upload full PowerPoint or PDF** – each slide becomes a separate scene in your video
- **Edit each scene individually** – add avatars, extra text, images, or video segments
- **Animate content** – apply motion effects, scene transitions, and timed pauses for smooth delivery
- **Integrate custom voice tracks** – upload .mp3 recordings (e.g., from ElevenLabs) for narration
- **Automatic transcription of audio** to text
- **Easily replace or modify scenes** – update single parts or entire video sections without redoing the whole project
- **One-click “Generate” option** – renders the final video and makes it ready for download
- **Unlimited creative control** – combine animation, multimedia, and AI avatars for professional results



## 4. HeyGen – AI Video Generated



Empa

Materials Science and Technology

Anonymize  
Your Data

Matthias Roesslein - Scientific IT

MP4 videos in full HD (1080p), with durations of up to 30 minutes per video.

4K resolution available – ideal for ultra-high quality projects, though costs may be prohibitive for private use.

# 4. Publish the Videos internally

- **Created in record time** – each video completed within **4 – 6 hours**
- **Open access** – all videos freely available on **Zenodo** for viewing and download
  1. Part: Introduction and data lifecycle
  2. Part: Data Management Plan (DMP)
  3. Part: Document your data
  4. Part: Format your data
  5. Part: Store and backup your data
  6. Part: Share your data in a project
  7. Part: Anonymize your data
  8. Part: Choose the appropriate copyright license
  9. Part: Publish your Data in Zenodo (Repository)

<https://doi.org/10.5281/zenodo.15878642>

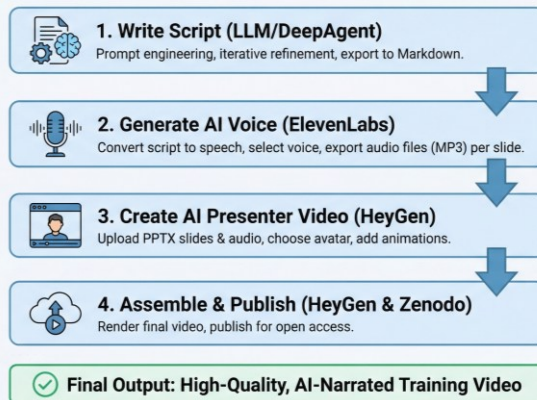
The screenshot displays the Zenodo profile of Matthias Roesslein. The interface includes a top navigation bar with the Zenodo logo, a search bar, and links to 'Communities' and 'My dashboard'. The user's name 'Matthias Roesslein' and email 'Empa' are shown. Below this, there are tabs for 'Uploads', 'Communities', and 'Requests'. The 'Uploads' tab is active, showing a list of uploads. Each upload entry includes a version number, date, and a title. The uploads are: 1. 'Composition data of cars with different drivetrains (ICE, BEV, HEV, PHEV, and Others) from 1980 to 2050' (October 14, 2025, Version 1.0.0, Dataset, Restricted). 2. 'Research Data Management - 9. Part: Publish your Data in Zenodo (Repository)' (September 7, 2025, Video/Audio, Open). 3. 'Research Data Management - 8. Part: Choose the appropriate Copyright License' (September 7, 2025, Video/Audio, Open). 4. 'Research Data Management - 6. Part: Share your data in a project' (August 23, 2025, Video/Audio, Open). 5. 'Research Data Management - 7. Part: Anonymize your data' (August 23, 2025, Video/Audio, Open). 6. 'Research Data Management - 5. Part: Store and backup your data' (July 27, 2025, Video/Audio, Open). Each entry also shows the number of views and downloads.

# Technology enabling such training videos



- **DeepAgent** – Provides access to a powerful AI agent plus ~32 LLMs for small tasks (e.g., text optimization, picture generation, ...). Automatically selects the best model for the job.
  - 💰 Cost: 20 US\$/month
- **ElevenLabs** – Used for all voice tracks, delivering high-quality voices and audio output.
  - 💰 Creator Plan: 18.33 US\$/month (annual subscription) — includes 100 min audio/month
- **HeyGen** – AI-powered avatar and video generation.
  - 💰 Creator Plan: 24 USD/month (annual subscription)
  - 💰 Business Plan: 119 US\$/month + 20 US\$/month additional seat

## Overview of the AI Training Video Creation Procedure



# New developments – AI generated scientific figures



**A strong figure is not just beautiful — it communicates one clear scientific message**

- Start with the question: **What should the audience understand in 5 seconds?**
- Choose the format: data plot, mechanism, workflow, or graphical abstract
- Separate **measured data** from **conceptual illustration**
- Use AI to visualize ideas — not to fabricate evidence

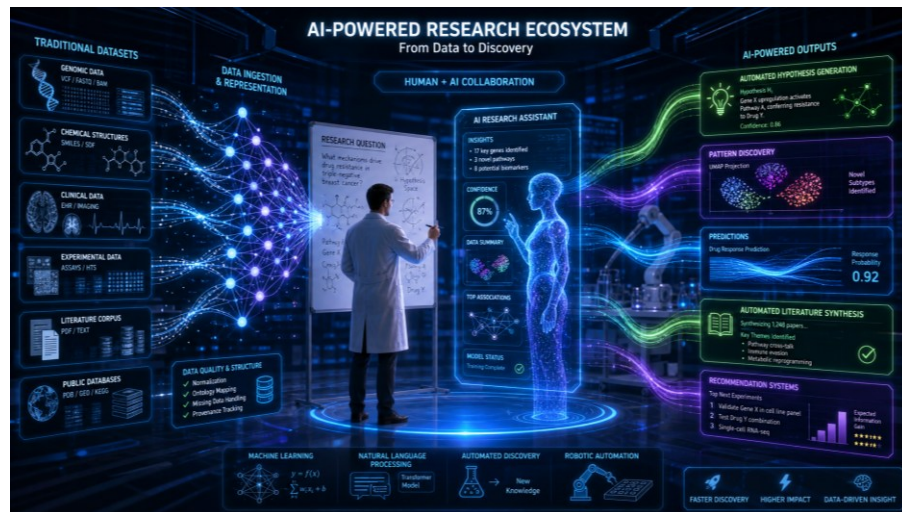


# Use the right tools for the job



**AI is powerful, but it is not the right tool for every part of a figure**

- **AI image models:** concepts, biological scenes, mechanisms, visual inspiration
- **Scientific software:** graphs, statistics, error bars, axes, and quantitative results
- **Design tools:** final layout, arrows, labels, and consistent styling
- **Editing AI:** targeted fixes, background removal, recoloring, and image expansion

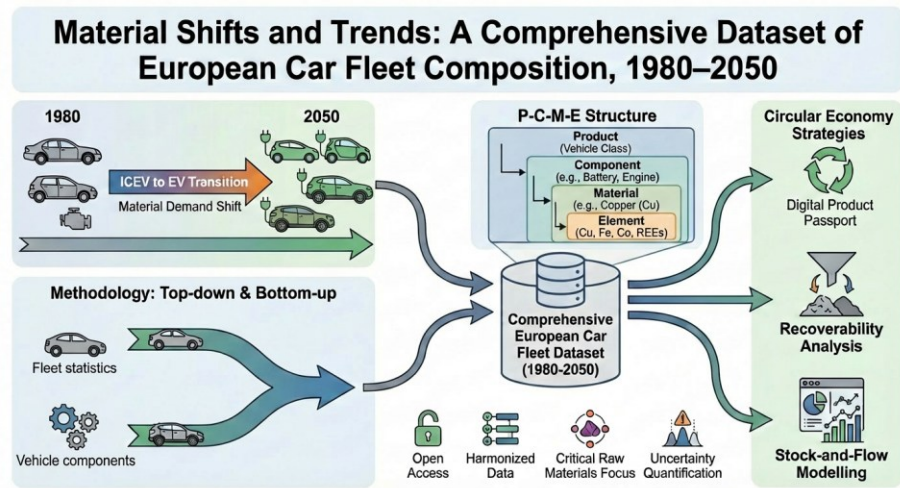


# Prompt like a figure designer



A vague prompt produces a vague image  
Specify:

- **What:** the scientific subject and key components
- **Why:** the message the figure must communicate
- **Where:** position and relationship of each element
- **How:** schematic, realistic, minimal, or journal-style
- **Avoid:** incorrect anatomy, decorative clutter, unreadable text, invented details



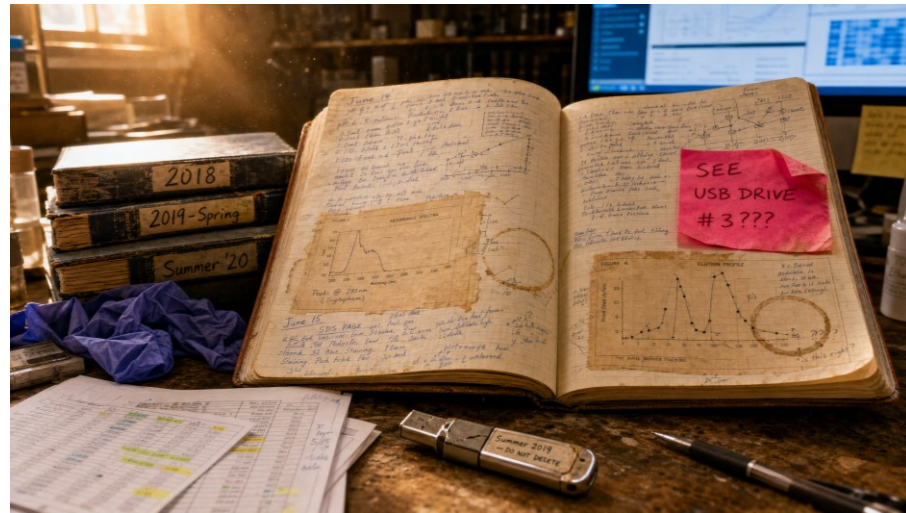
# AI can look right — and still be wrong



**Scientific plausibility must be checked, not assumed**

**Watch for:**

- Wrong structures, shapes, or experimental setup
- Arrows pointing in misleading directions
- Fake labels, spelling errors, and unreadable text
- Incorrect scale, proportions, or number of components
- Attractive details that have no scientific basis



# Scientific motion: The power of AI video



## Modern video models bring static concepts into dynamic, cinematic motion

- **Bring processes to life:** Visualize fluid dynamics, cellular uptake, drug delivery, or mechanical workflows in motion
- **Cinematic storytelling:** Dynamic camera angles, depth of field, and continuous motion that capture attention instantly
- **Ideal formats:** 5 to 15-second visual hooks for conference presentations, web teasers, and graphical abstract animations
- **Idea exploration:** Quickly test how a complex 3D mechanism or physical interaction could look before commissioning custom animation

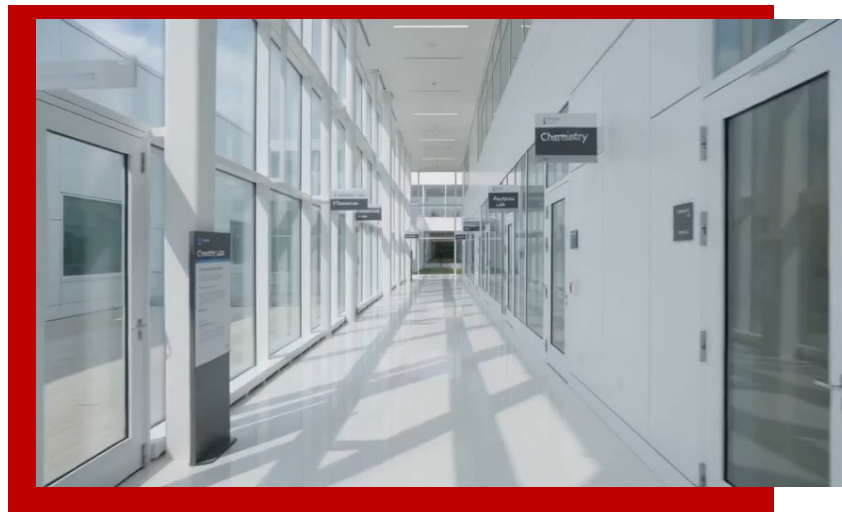


# Where AI video falls short: The scientific reality



## Visual realism does not equal physical or biological accuracy

- **Hallucinated mechanics:** Objects can morph unnaturally, merge into one another, or violate laws of physics
- **No scientific precision:** The model does not understand real binding kinetics, stoichiometric ratios, or exact molecular geometry
- **Text & labels are unusable:** On-screen text, quantitative meters, and scale bars will almost always be scrambled or distorted
- **Consistency drift:** Maintaining the exact same protein shape or device across multiple camera cuts is extremely difficult



# The hybrid workflow: From raw clip to final scene



**High-impact scientific video is built in layers—AI creates the backdrop, you supply the facts**

- **Step 1 - Generate the core motion:** Create clean, slow-paced visual footage focusing purely on background and movement
- **Step 2 - Overlay clean graphics:** Add accurate callout lines, annotations, molecular formulas, and legends in video editing software
- **Step 3 - Insert verified data:** Composite real plots, high-res static structures, or microscope overlays directly onto the video
- **Step 4 - Control the narrative:** Use voiceover, pacing, and clear title cards to anchor the scientific context



# Best practices: Getting the most out of video prompts



**Direct the camera, define the motion, and keep the composition clean.**

- **Describe camera movement explicitly:** Specify slow cinematic zoom-in, orbital pan, or microscopic tracking shot.
- **Keep movement simple:** Focus on one clear physical process per clip rather than chaotic multiple events
- **Prompt for clean framing:** Ask for uncluttered backgrounds and clear focal subjects to leave room for future text overlays
- **Generate short, combine later:** Produce multiple 5–10 second focused clips and cut them together in an editor rather than asking for one long complex sequence



# Creativity has no limits, ...



- **Market Saturation:** An "inflation" of AI tools, making professional video generation accessible to everyone
- **Creative Speed:** With minimal effort, these tools allow for the rapid creation of high-quality, imaginative stories
- **Platform Longevity:** Be cautious when choosing tools, as many new platforms may not remain available
- **Data Sovereignty:** Verify where these platforms are hosted and how they manage sensitive biometric data
- **Security Compliance:** Prioritize tools that are officially certified to meet global security standards like SOC 2, GDPR, and HIPAA



# EU AI Act, Art. 50 – what it means for our training videos



| Element used          | Obligation                                                                                         | What we do concretely                                                                                    |
|-----------------------|----------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|
| AI voice narration    | Synthetic audio must be disclosed; cloning a real person's voice requires their documented consent | Use a licensed stock voice, state "narration generated with text-to-speech" in the intro and description |
| AI avatar / presenter | Clearest case: an avatar resembling a real person is treated as a deepfake and must be labelled    | Visible on-screen notice in the first seconds, plus a permanent corner label or lower third              |
| AI-generated images   | Providers must embed machine-readable marking; we may not remove or obscure it                     | Keep C2PA metadata on export, note generated illustrations in the credits                                |
| AI video sequences    | Same as images; applies to B-roll, animations and any synthetically altered footage                | Check that rendering and re-encoding preserve watermarks and metadata                                    |
| AI-assisted script    | Human review and editorial responsibility reduce the obligation, but do not remove good practice   | Named human author reviews and signs off every version before publication                                |
| Whole production      | Disclosure must be clear, distinguishable and given at first exposure; non-compliance is fineable  | Write the disclosure into the script itself, so it survives every regeneration; document the toolchain   |

# Training videos



Matthias Rösslein

[matthias.roesslein@empa.ch](mailto:matthias.roesslein@empa.ch)

A graphic with a cream-colored background and abstract blue and red shapes in the corners. A thick red L-shaped border is on the left and bottom. The text "THANK YOU FOR YOUR ATTENTION" is centered in dark blue.

**THANK YOU  
FOR YOUR  
ATTENTION**