The Alchemy of AI Generation: Samplers, Schedulers, and Creative Settings

Unleashing the Machine Muse (Introduction)

Generative AI can feel like digital magic – type a prompt and *poof*, out comes a vivid image, a musical clip, or even a short video scene. But behind this apparent sorcery are a bunch of **dials and knobs** that control *how* the AI creates content. These include mysterious-sounding things like **samplers**, **schedulers**, **steps**, **noise levels**, **temperature**, and **guidance scale**. Think of these as the settings on an "AI creativity console." The goal of this guide is to demystify these concepts in **image, audio, and video generation**. We'll explain what each does, how they affect the final output, and draw parallels across different media. By the end, you'll see how an AI artist or composer isn't all that different from a gamer tweaking settings for the perfect experience – it's witty, a bit technical, but ultimately empowering and fun.

Samplers and Schedulers: Orchestrating the Creative Process

In generative AI (especially **diffusion models** for images), *samplers* and *schedulers* refer to the algorithms that **orchestrate the step-by-step creation** of content. In fact, the terms are often used interchangeably. A scheduler (or sampler) is basically the **game plan for generation –** it guides how the AI transforms pure randomness into coherent output . Specifically, it determines:

1. **Number and size of steps:** How many iterative steps the process will take, and how big each step is (e.g. big leaps at first, then fine tweaks towards the end) .
2. **Stochastic vs. deterministic behavior:** Whether the process injects randomness at each step or follows a fixed trajectory . In other words, *will the result be a unique surprise each time or reproducible if we use the same seed?*

3. **The algorithmic method of generation:** The mathematical strategy used to refine the output (different algorithms approach the task in their own style) .

In plainer terms, the sampler is like the **director of a movie**, telling the AI *how* to go from static noise to a final masterpiece. Different directors have different styles – and so do different samplers. For example, image generators offer a menu of samplers such as **Euler, Euler Ancestral, DDIM, PLMS,** or **DPM**. Each comes with its quirks and trade-offs:

- *Euler* (and its "ancestral" variant) is known for speed, often needing fewer steps for good results . It's the quick sketch artist – great for rapid concept drafts.
- *DDIM* (Denoising Diffusion Implicit Model) was one of the first diffusion samplers, designed for efficiency and faster processing .
- *DPM solvers* (and newer **DPM++** versions) use fancy differential equation math to improve image quality, with options for one-step or multistep refinement . Think of these as detail-oriented painters – possibly slower, but aiming for high fidelity.
- *Ancestral samplers* (often denoted by an **"a"** in their name, like Euler *a*) add extra noise at each step, introducing more randomness . They behave like an improv musician riffing on a theme – each performance (image generation) comes out slightly different, even with the same starting conditions. Non-ancestral (deterministic) samplers, by contrast, will converge to the **same result given the same seed**, acting more like a classical musician playing a score note-for-note .

Why Sampler Choice Matters

Choosing a sampler/scheduler in image generation can **significantly influence the final output** . It's not just academic – it affects the *look, feel, and quality* of the content:

- **Quality vs. Speed:** Some samplers produce higher quality images but take more steps (slower), while others are faster but might miss fine details. It's a classic *performance vs. quality* slider – e.g., Euler is quick and often good enough, whereas DPM2 Karras (a variant with a special noise schedule) might squeeze out extra detail if you're willing to wait a bit .
- **Style and Detail:** Different algorithms can have subtle effects on style. One sampler might yield softer, smoother outputs; another might emphasize

textures or contrast. Under the hood they all follow the same overall diffusion recipe, but their "brushstrokes" differ. For instance, one might take larger early steps to nail the composition and smaller late steps for details, versus another that changes step sizes differently – influencing whether the result is more *globally coherent vs. locally detailed* .

- **Determinism vs. Variety:** As noted, using a deterministic sampler with a fixed random seed will give you the same result every time – great for reproducibility (vital if you need consistency, say for a game asset). In contrast, a stochastic sampler (ancestral) injects a bit of randomness each iteration, so you get a fresh variation with each run . This is like asking the AI to "try it again, but different this time." It's useful when you want to generate many variants of a concept (like multiple concept art sketches) without changing the prompt.

In audio and video generation, the notion of a *sampler* is present too, though it may look different. For example, in an AI music generator, the "sampler" might be the strategy used to pick the next note or sound chunk from a probability distribution (more on *temperature* for this soon). And in AI video, which often builds on image diffusion tech, you also have **schedulers** controlling the frame-by-frame denoising process – essentially the same role as image samplers but extended to 3D (2D frames over time).

The key takeaway: **Samplers/schedulers are your creative AI's game plan**. Changing that plan can mean the difference between a quick rough draft or a polished piece, a consistent reproduction or a one-of-a-kind surprise. Don't be afraid to experiment – as one guide cheekily put it, *"TL;DR: play around with different schedulers once you're happy with an image to get to the final image!"* . The AI won't complain – it actually loves trying new methods at your command.

Steps and Noise: From Chaos to Creation

Most generative models start from *chaos* – quite literally in the case of diffusion models, which begin with a blob of random noise. The process of generation is essentially **taming that noise over a series of steps**. The **number of steps** is like the number of iterations or brush strokes the AI will use to refine the output. More steps generally allow finer detail (imagine an artist who takes 100 strokes to refine a painting versus one who stops at 20 strokes) – but with diminishing returns and more computation time . Fewer steps make generation faster but can leave the result looking undercooked or distorted, as the AI might not have had enough rounds to polish the output .

Each step in a diffusion model reduces the noise slightly, gradually revealing the image (or audio/video) content. The **noise schedule** defines how this reduction happens – typically a high noise level at the start, tapering off to zero noise at the final step . Early steps make broad changes (laying out composition or melody), and later steps add fine details (refining textures or musical flourishes) . At each step, the *sampler* ensures the generated content matches the noise level dictated by this schedule .

Illustration: An AI image generation process starting from pure noise and becoming clearer with each denoising step . The sampler here gradually removes noise to reveal the content. Early on it's just static; by the final step, a recognizable image emerges. (In practice this was done with the Euler sampler over 15 steps – notice how each frame gets less grainy and more detailed.)

The **initial noise** is determined by a *random seed*. If you use the same seed and the same settings, a deterministic sampler will produce the *exact same output* – great for when you need to recreate that perfect image later. Change the seed, and you get a *"parallel universe"* version of the image: similar overall look, but with various details changed . As one source nicely described, different seeds make the diffusion process take a "slightly different trajectory but end at a similar place" – like alternate takes on the same idea. On the other hand, changing the **prompt** (the textual or conditional input) sends the process to a *totally different destination* – you'll get a drastically different result altogether .

In **audio** diffusion models (yes, they exist!), the idea is analogous: the model might start with a burst of noise that sounds like radio static and then refine it into, say, music or speech over many steps. The concept of steps and noise applies one-to-one. For example, "Stable Audio" (by Stability AI) and others use latent diffusion on audio spectrograms – beginning with noisy spectrograms and cleaning them into music. If the model is *not* diffusion-based (some audio models are autoregressive, generating one chunk after another), the "steps" correspond to each note or sound generated sequentially. There may not be an initial noise in that case, but there's still an initial state (often random) from which generation begins.

The bottom line: **steps and noise are the canvas and raw material of generative art.** More steps = more chances to refine; initial noise = the seed of creativity (or randomness). In practice, you'll often tweak the number of steps to balance quality vs. speed. And if you want a new variation, you can just roll a new random seed (the AI equivalent of *shuffling the deck* before dealing a new hand).

Temperature: Tuning the Chaos

If samplers are the *plan* and steps are the *effort*, **temperature** is the setting that controls the *wildness* of the AI's imagination. This concept comes up in many generative models, especially those that produce sequences (text, music, etc.), but it's broadly applicable. **Temperature** is basically **randomness control**. Technically, it skews how the AI samples from the probability distribution of possible outputs. In human terms, it's like telling the AI to **"keep it conservative"** versus **"get creative, don't be shy."**

- **Low temperature (e.g. 0.2 or 0.5):** The AI becomes more *deterministic*. It will pick the most likely or sensible continuation each time. The output is more predictable and coherent – no curveballs . This is like an improv comedian *playing it safe* and sticking to well-trodden jokes.
- **High temperature (e.g. 1.0, 1.5 or beyond):** The AI will take more risks and throw in more randomness. Outputs become more diverse, surprising, sometimes off-the-wall. You get more novelty, but also a higher chance of nonsense or errors . This is our comedian going wild with experimental, bizarre humor – might flop or might be brilliant. In AI music, for instance, a higher temperature might introduce unusual riffs or instrument choices that weren't the most obvious next note, adding variety .

In practical terms, *temperature* is a dial you often see in AI text generators (like GPT-based systems) and AI music generators. For example, Meta's **AudioCraft** (which includes music generation models) lets users set a temperature when sampling notes. A higher temperature yields more diverse or unexpected musical elements . The AWS blog on AudioCraft notes: *"A higher temperature increases the randomness of the output, making it more diverse."* If you keep temperature at 0 (or very low), the model will output its best guess every time – that can converge on a single answer (great for factual text, or a consistent melody), but may lack creativity or variation.

In image diffusion tools, you might not see an explicit "temperature" slider, because diffusion randomness is usually controlled by the **random seed and sampler choice** (ancestral samplers implicitly add randomness akin to a higher temperature). However, some image generation setups allow an eta parameter (for DDIM sampler) or similar, which essentially reintroduces some randomness into an otherwise deterministic process – effectively functioning like temperature. The concept is the same: **low randomness for predictable refinement, high randomness for more divergent outcomes**.

Analogy: If generating content is like making soup, temperature is how high you turn up the heat under the cauldron of creativity. Low heat simmers slowly and safely; high heat might boil over with wild new flavors. Neither is "correct" – it depends if you want a safe recipe or a surprise stew. Finding the right temperature can be crucial: many creative users find a sweet spot where the output is **interesting but not gibberish**. For instance, in text generation a temperature around 0.7 is often a good mix of coherence and creativity. In music, one might start with temperature ~1.0 for some randomness and adjust if the output is too boring or too chaotic.

Guidance Scale: How Clingy is Your AI Muse?

Moving on to **guidance scale** – this setting is specific to *conditional generation* (when you guide the AI with an input like a text prompt). In image generation (Stable Diffusion and the like), it's usually called **CFG (Classifier-Free Guidance) scale**. In simple terms, **guidance scale controls how strictly the AI follows your prompt or instructions** . It's the *obedience knob* for the AI: low values mean the AI has more freedom to improvise, high values mean it's handcuffed to your prompt (for better or worse).

How does it work? Under the hood, the model actually generates two predictions – one with the prompt and one without (unconditioned) – and guidance scale blends them. But you don't need to know the math to use it:

- **Low guidance (close to 1 or 2):** The model only *loosely* follows the prompt. It might drift or come up with something only tangentially related. This can yield more *creative or unexpected* results, but sometimes it just ignores important parts of your request . Imagine telling an AI "a cat sitting on a beach" with low guidance – you might get a weird scenery where the cat is barely visible or it's something like a tiger in a jungle because the model free-associated.
- **High guidance (e.g. 15 or 20 in Stable Diffusion):** The model *strictly* adheres to the prompt. If you say "a cat sitting on a beach," it will *really try* to put that cat on that beach, exactly as asked. Sounds ideal, but there's a catch: pushing guidance too high can **reduce output quality** or diversity . The image might look unnatural or have artifacts because the model is

overexerting itself to satisfy every word. It's like an employee doing *exactly* what the boss says even if it makes the project worse. In practice, extremely high guidance can make images look brittle or weird, and in music, a high guidance might force the melody to stick to a motif so hard that it sounds strained .

Typically, there's a **sweet spot**. For Stable Diffusion, many users find that a guidance scale around 7–12 gives a good balance – the image closely matches the prompt without too many artifacts . At guidance = 1 (effectively no guidance), the prompt might be largely ignored; at guidance = 20, the image follows every word but often looks *worse* overall . One interactive guide shows that at guidance 7 you get a nicely composed "panda playing guitar" as prompted, while at 1 the panda might be missing, and at 20 you get a very literal but garbled panda-guitar chimera .

This concept extends to audio and video as well. If you're using a text-to-music model (like MusicGen in AudioCraft), there is a guidance scale parameter that does the same thing: it weighs how much the model should listen to the text prompt's "desires" versus its own learned musical distribution. The AudioCraft team suggests that **guidance above 1.0** makes the output follow the prompt more closely but can harm audio quality, and they found *around 3* to be a good default for music . In fact, *"a higher guidance scale encourages the model to generate samples more closely linked to the input prompt, usually at the expense of poorer audio quality"* . Video diffusion models (like those that take text and generate a short video) also use guidance scale in the same way – e.g., one text-to-video pipeline uses a default guidance ~7.5, and notes that higher values make the video scenes stick to the prompt but can lower the video quality .

Think of guidance like a slider between *"Artist, do your thing"* and *"No, do it exactly like I imagined!"*. Neither extreme is inherently bad: low guidance can spark *happy accidents* and originality, while high guidance ensures *relevance* to your instructions. For creative workflows, one strategy is to start with a moderate guidance (so the output is on-topic), then adjust if needed – if the output is too generic or off-prompt, nudge guidance up; if it's too on-the-nose and lacks creativity, ease guidance down. It's almost like collaborating with the AI: how much freedom do you give your **robot artist** or **AI composer**?

Different Mediums, Same Ideas: Image vs. Audio vs. Video

Now that we've covered these key concepts, it's worth highlighting how they **translate across image, audio, and video generation**. While the underlying AI models for each medium can be very different (pixel diffusion vs. waveform synthesis vs. frame interpolation), the *conceptual knobs* we've described have surprisingly similar roles in each domain.

- **Image Generation:** Typically uses diffusion models (e.g. Stable Diffusion) or other generative networks (GANs, etc.). Here, *samplers/schedulers* are front and center – you pick one to generate your image, and you decide on the number of *steps*. You tweak *guidance scale* to balance prompt fidelity vs. creativity, and the *random seed/noise* gives you different variations. *Temperature* isn't explicitly adjusted in most diffusion-based image tools, but if you use an image GAN or a generative model that supports it, temperature could control diversity of outputs. For example, some image models allow variation in sampling for diversity (like the "variation seed" in Stable Diffusion which adds a subtle noise offset).
- **Audio Generation:** There are a few flavors of AI audio. **Diffusion-based audio** (like vocoders or Stable Audio) will mirror the image approach: they start with noise and have a sampler, steps, etc., plus maybe a guidance scale if conditioned on text. **Autoregressive or transformer-based music models** (like OpenAI's Jukebox, Google's MusicLM, or Meta's AudioCraft/MusicGen) generate audio as a sequence of tokens or samples. In those, you won't hear about "denoising steps," but you'll see *temperature*, *top-k*, *top-p* (nucleus sampling) and similar parameters for how to sample each next chunk. The concept of *guidance* still appears if the model is conditioned on something (text prompt, or a melody to follow). Indeed, MusicGen uses classifier-free guidance under the hood (with a default guidance scale of 3 for text conditioning) to better align the music with the prompt . So an audio model user might say: "I set the **temperature high** to make the music more unpredictable, but kept **guidance** moderate so it still feels like the 'jazzy hip-hop' I asked for." The parallels to image generation are clear – randomness vs. fidelity.
- **Video Generation:** This is the new frontier, but many text-to-video or image-to-video models are basically extensions of image diffusion. They generate either a series of images with a temporal coherence mechanism, or directly generate a 3D spatial-temporal block of data. For example, ModelScope's text-to-video diffusion and Runway's Gen-2 take the Stable Diffusion idea and add a time dimension. Consequently, you'll find **familiar**

parameters: number of inference *steps*, a *scheduler* choice (e.g. DDIM vs. others, just like images) , and *guidance scale* (to follow the text prompt) . One difference is that videos are heavier to compute, so often models use fewer steps or lower resolution, and there's an extra challenge of keeping things consistent across frames (no one wants a character that shapeshifts every frame!). Some video models use *two types of noise*: one that is global (same for all frames to maintain coherence) and one that is frame-specific . But as a user, you still largely control it with the same knobs – e.g. "I used 50 steps with a DDIM sampler to generate a 2-second video, and a guidance scale of 8 to make sure the scene stayed true to 'Darth Vader surfing a wave' (yes, that's possible now) ."

In summary, **the language of generation settings translates well between domains**. An artist using Stable Diffusion for images, a musician using AudioCraft for tunes, and a filmmaker dabbling with Gen-2 for videos will all find themselves adjusting how *precise vs. exploratory* the AI is (guidance), how *much randomness* to allow (temperature/noise), and how *long or iterative* the generation should be (steps). The major differences lie in the data being produced – pixels vs. sound waves vs. moving frames – and the model architectures. But conceptually, these settings serve analogous purposes. Once you grasp them in one domain, you've got a head start in another. For instance, an undergrad who played with text prompts and CFG scale in an image generator will intuitively understand why a text-to-music model might need a guidance scale to decide how closely to follow the mood described in words.

One more cross-domain note: **reproducibility** via seeds applies to any generative model that is stochastic. If you can set a seed in an audio model or video model, doing so will let you regenerate the same output (or at least very close) later, as long as other parameters are the same. This can be crucial for creative workflows – imagine a game that uses AI to generate music on the fly; you might want to be able to replay that exact music later, so you'd keep the seed. Or a VFX artist generating an AI video may lock a seed once they get a great result so that they can upscale or refine it consistently. **Seeds are your friend** for controlling randomness across all these mediums.

Creative Applications in Games and Media (Use Cases)

So far we've been knee-deep in the *how* of generative AI. Let's zoom out and look at the *what for*. Here are some **real-world use cases** in creative fields and gaming that show these concepts in action:

- **Concept Art Generation:** Generative AI is a godsend for concept artists and game designers in the brainstorming phase. With tools like Stable Diffusion, Midjourney, or others, you can whip up dozens of visual ideas in minutes – from character designs to environment art – by tweaking prompts and sampler settings. This allows rapid iteration and exploration of styles that might not emerge via traditional sketching. In practice, an artist might generate a batch of images with a prompt like "ancient forest temple, overgrown with roots, mystical lighting" and use an **ancestral sampler** to get varied outputs each time. According to one report, AI has made *"significant strides in concept art, allowing artists to quickly generate and iterate on ideas."* It not only **speeds up the creative process** but also *"enables the exploration of new artistic styles and ideas that might not have been conceived by human artists alone."* Game studios can use this for mood boards and inspiration, saving time and giving human artists a springboard to refine further.
- **AI Music Scoring:** Composers for film, TV, and games are experimenting with AI to generate musical backdrops and even complete scores. Imagine being a game developer who needs an ambient track for a level – instead of trawling through stock music or hiring a composer for every tweak, you could use an AI music generator to produce a fitting score, then adjust settings to suit the mood. For example, you might set a **low temperature** for a predictable, steady background drone in a horror game (no surprise trumpet solos, please), or a higher temperature for an unpredictable, evolving soundscape in a sci-fi scene. AI tools like **AIVA** are explicitly used for this: AIVA is an AI composer that can *"generate background scores for films and TV shows"* and help create *"immersive soundtracks for video games."* The guidance-like controls let you specify style (e.g. "suspenseful orchestral") and the AI sticks to it. This doesn't mean composers are out of a job – rather, they can use AI to draft pieces and then fine-tune or build on them. It's akin to having a tireless assistant who can jam 24/7. We're already seeing AI-generated music in indie games and as temp tracks in filmmaking. As the tech improves, dynamic game scoring (where the music adapts to gameplay in real-time) becomes more feasible. In fact, some modern game engines could use an AI model to tweak the music on the fly

(ramping up intensity during a boss fight by increasing the *"drama" prompt weight or guidance*).

- **Cinematic Animation (AI-Generated Video):** Generative AI is pushing into animation and video production, enabling short "AI films" or VFX sequences made with minimal crew. A sharp-toned example: filmmakers have used tools like Runway **Gen-2** (text-to-video) in combination with Stable Diffusion to create entire short films. One filmmaker, Paul Trillo, generated storyboard images with Stable Diffusion, then fed them and a prompt into Gen-2 to get moving clips . The process was iterative – he describes it as *"shaking a magic eight-ball until you get the answer you're looking for"* , because you often have to run the generation many times (changing seeds, prompts, or using negative prompts to fix weird artifacts) to get a usable result. The current state-of-the-art can produce surreal, dreamlike visuals that would be hard to film or animate manually. Trillo's film *Thank You For Not Answering* is one example of a 100% AI-generated short – showcasing how these tools can be used like an "avant-garde paintbrush" to achieve an aesthetic that's *"otherworldly, like a snapshot of a dream"* . In practice, an animator using AI might set **high guidance** to make sure the video sticks to the script of the scene, but might use a variety of seeds and even different samplers to generate alternate takes of the scene. As this tech matures, we could see it used for pre-visualization in filmmaking (quickly animating a storyboard to see how a scene might play out), or stylized content like music videos and concept trailers. It's worth noting that because video models are still finicky, creators often have to *curate or stitch together* outputs (like Trillo generating 400 clips to make a 5-minute film) – so the sampler settings and prompt engineering are part of a new animation workflow.
- **Dynamic Asset Creation in Games:** Perhaps one of the most exciting use cases is using generative AI for **dynamic, real-time content generation** in games. This can range from on-the-fly generation of NPC dialogue, to creating endless procedural worlds, to generating textures or 3D models within a game as players progress. The idea is that a game could feel truly limitless and adaptive. For example, consider a role-playing game where every town's layout and artwork is freshly generated, or a survival game where the soundtrack and sound effects adapt to what's happening to the player. Generative AI can power this by taking player state or actions as input (prompts) and producing content immediately. A study on AI in

gaming highlighted *"dynamic generation of in-game elements in real-time – dialogues, terrain, soundscapes, and music that adapt to the player's actions and decisions."* This means no two playthroughs are the same . From a settings perspective, this might involve keeping *temperature somewhat high* to ensure variety (you don't want every forest to look identical), but using *guidance* or constraints to ensure the generated content fits the game's theme and previously generated assets. For instance, if the game's art style is set by earlier content, one could guide the AI art generator to stick to that style (some advanced pipelines use *image prompts or fine-tuned models* for consistency). Similarly, music could be generated to match the mood: peaceful melody in safe areas vs. tense rhythm in combat, triggered by the game and guided via prompt. We already see early signs of this: **procedural level generation** with AI, NPC dialogue generation (e.g., AI dungeon masters in tabletop-style games generating narrative on the fly), and adaptive soundtrack engines. As one gaming blog put it, these dynamic AI generation capabilities *"enhance player immersion and replayability, offering a unique experience every time."* It's like having a game that writes itself as you play, within the boundaries the designers set.

Finally, while not explicitly in the list, it's worth mentioning **other creative tech** intersections: for example, AI-driven **voice generation** for games (creating NPC voices with TTS on the fly), which uses models with similar knobs (you might adjust speaking *style* as a kind of guidance). Or **3D asset generation** (some tools attempt to generate 3D models or textures from text) where concepts like *diffusion steps* and *guidance* are used in generating 3D geometry or mapping images. The same principles we covered echo here – controlling fidelity to prompt vs. randomness is key to getting useful results that still have variety.

Wrapping up: Understanding samplers, schedulers, and other generation settings is like getting the "cheat codes" to your AI creative tools. Rather than treating the AI as a mysterious oracle, you can actively **steer the output**. Need a quick idea dump? Use a fast sampler with few steps and low guidance to let the AI run wild. Need a polished result? Crank up the steps, pick a precise sampler, maybe raise guidance a bit to stick to your vision, and lower temperature to avoid chaos. Across images, music, and video, these dials let you decide whether the AI should behave like a strict rule-following student or a jazz artist improvising on a theme. And for those in gaming and interactive media, mastering these settings opens up new frontiers of **real-time creativity**, where games can generate content on the fly that feels both fresh and context-appropriate.

Ultimately, generative AI is **part science, part art** – and knowing these key concepts turns you from a passive consumer into an active artist or designer collaborating with the machine. So go ahead, experiment with that sampler or scheduler you haven't tried yet, play with the guidance scale, and don't fear the noise – within that randomness lies infinite potential for innovation in creative tech and gaming.

1. DiffusionTools. (2023). *Samplers and Schedulers in Stable Diffusion*. https://diffusiontools.gitbook.io/stable-diffusion/samplers-and-schedulers
2. Hugging Face. (2023). *What are classifier-free guidance scale and inference steps?*. https://huggingface.co/blog/stable_diffusion#what-are-classifier-free-guidance-scale-and-inference-steps
3. AWS Machine Learning Blog. (2023). *Generate music using AudioCraft and Hugging Face on Amazon SageMaker*. https://aws.amazon.com/blogs/machine-learning/generate-music-using-audiocraft-and-hugging-face-on-amazon-sagemaker/
4. Trillo, P. (2023). *Thank You For Not Answering [Short film]*. https://www.runwayml.com/blog/thank-you-for-not-answering/
5. GameDev Academy. (2023). *How generative AI is reshaping video game design*. https://gamedevacademy.org/generative-ai-in-game-design/
6. AIVA. (2023). *AIVA – The AI composer that creates musical pieces*. https://www.aiva.ai/