Generative models

Complete Guide to Free, Locally-Runnable Generative Al Models

This guide consolidates all information from both sources into a single, structured reference for free, opensource generative AI models that can be run entirely locally on your own hardware. It is organized by output type, with clear categorization, model details, capabilities, hardware requirements, workflows, and integration paths for app development.

All listed tools are open-source, free for personal and often commercial use, and designed to operate offline -ensuring privacy, no API costs, and full control over your data.

Key Considerations Before You Start

Hardware: A capable NVIDIA GPU with at least 8GB of VRAM is recommended for most tasks, especially image and video generation. Some lightweight audio and image models can run on CPUs or GPUs with 4GB VRAM.

Storage: Model files are large, typically ranging from 2 GB to 12 GB each. Plan your storage accordingly (50-100 GB free space recommended).

Software: Most models require a Python environment with deep learning libraries like PyTorch or TensorFlow. User-friendly frameworks like ComfyUl, AUTOMATIC1111, or GPT4All simplify this process.

Workflow: Generated assets (especially UI mockups and vectors) often require post-processing and polishing in standard design tools like Figma, Penpot, or Inkscape.

Privacy & Cost: Local execution ensures no data leaves your machine and no API fees.



Core Frameworks & Platforms for Local Execution

Before diving into specific models, you need a local environment to run them. These platforms simplify installation, management, and orchestration of multiple Al models.

Tool	Capabilities	Hardware Requirements	Notes
ComfyUl	Node-based visual workflow for Stable Diffusion, AnimateDiff, audio, video. Highly customizable.	GPU recommended (NVIDIA 8GB+ VRAM ideal)	Best for advanced users; supports complex pipelines.
AUTOMATIC1111 WebUI	Popular web interface for Stable Diffusion. Rich plugin ecosystem (ControlNet, LoRA, T2I-Adapter).	4–8 GB VRAM	Great for beginners and pros alike.

Tool	Capabilities	Hardware Requirements	Notes
InvokeAl	Professional-grade UI with canvas editing, inpainting, outpainting, and layer support.	4+ GB VRAM	Ideal for iterative UI asset creation.
Fooocus	Simplified, one-click Stable Diffusion frontend with preset styles and aspect ratios.	6+ GB VRAM	Minimal setup; excellent for app icons and mockups.
LocalAl	OpenAl-compatible REST API for running LLMs, image, audio, and experimental video models locally.	CPU/GPU (varies)	Unified API; supports agents (LocalAGI), semantic search (LocalRecall).
Ollama	CLI-based model runner with growing support for multimodal models. Easy to script.	CPU/GPU (optional)	Lightweight; integrates well with Penpot, Figma via plugins.
GPT4AII	User-friendly desktop app for LLMs and image models. No GPU required.	4–8 GB RAM	Beginner-friendly; good for prototyping.

Tip: Use conda or pip with PyTorch + CUDA for GPU acceleration. For NVIDIA GPUs, install CUDA 12.1+ and cuDNN. Mac users: M1/M2/M3 chips can run many models via MLX or DiffusionBee (macOS-specific).

1. Raster Image Generation (PNG, JPG) – App Graphics, Icons, UI Elements

Generate high-quality PNG assets for apps: icons, backgrounds, buttons, dashboards, avatars, etc.

▼ Top Models

Model	VRAM Req	Key Features	Notes
Stable Diffusion 1.5	4–6 GB	Foundational model; vast ecosystem of fine-tunes (LoRAs), ControlNet, IP-Adapter.	Use with Realistic Vision, DreamShaper, or OpenJourney for style-specific outputs.
Stable Diffusion XL (SDXL) 1.0	8–10 GB	Higher resolution, better prompt understanding, improved composition.	Preferred for modern UI generation.
SDXL- Lightning/Turbo	6–8 GB	4–8 steps, near-instant generation (~0.5 sec per image).	Great for rapid iteration.
Flux (flux.1- dev)	8–12 GB	Superior prompt adherence, handles text-in-image, hands, and details better than SDXL.	Emerging top-tier open-source alternative.

Model	VRAM Req	Key Features	Notes
DeepFloyd IF	12+ GB	Multi-stage diffusion for ultra-high-fidelity text rendering and complex prompts.	Best for Uls with embedded text.
AnimateDiff	8+ GB	Adds motion to static SD/SDXL images → animated sequences (GIF, MP4, Lottie).	Integrates with ComfyUI/AUTOMATIC1111.
GFPGAN	4+ GB	Face restoration and enhancement.	Fixes blurry or low-res avatars. Post-process portrait images.
AnimeGANv2	4+ GB	Converts photos to anime-style art.	Outputs clean line art suitable for vectorization. Lightweight; good for stylized assets.

Recommended Workflows

- Icons & UI Elements: Use SDXL + Fooocus with prompts like "flat design red notification bell icon, 512x512, white background".
- Mockups: Use Stable Diffusion with ControlNet (canny/depth) to generate consistent app screens.
- Photo Enhancement: Run GFPGAN after generating user avatars.

Installation Example

pip install diffusers torch transformers accelerate huggingface-cli
huggingface-cli download runwayml/stable-diffusion-xl-base-1.0 --local-dir
./models/sdxl

2. Vector Graphics (SVG) – Scalable Assets for Web & Apps

True Al-to-SVG models are still experimental. Most workflows involve raster-to-vector conversion.

✓ Available Tools & Workflows

Tool	Output	Requirements	Notes
SVG-LLAMA / SVG Diffusion	Direct SVG generation from text	Python + ComfyUI nodes	Experimental; limited availability.
DigitalMagicWand Al SVG Creator	Clean, laser-cuttable SVGs from prompts	Local Python setup	Open-source; suitable for icons and logos.
Stable Diffusion + ControlNet (Line Art)	High-contrast line art → vectorizable PNG	SD + ControlNet edge detection	Generate clean outlines, then trace.
Inkscape (Auto-Trace / Live Trace)	Bitmap → SVG	Free desktop app	Use with Potrace backend for best results.

Tool	Output	Requirements	Notes
VectorFusion	Diffusion-based raster-	1.4 GB model	Combines diffusion with
vectorrusion	to-SVG conversion	1.4 GB IIIOGEI	traditional tracing.

Recommended Workflow

- 1. Generate a high-contrast line drawing using SD + ControlNet (Canny or Scribble).
- 2. Export as PNG.
- 3. Import into Inkscape \rightarrow Path > Trace Bitmap \rightarrow export as SVG.
- 4. Use in Figma, Penpot, or export as Lottie for apps.

Installation

• **Inkscape**: Free, open-source, cross-platform vector editor.

3. UI & UX Design – Figma, Penpot, and Design System Integration

No model directly outputs .fig or .penpot files yet. However, several tools generate UI mockups or assist in code export.

▼ Tools & Plugins

Tool	Platform	Capabilities	Notes
Penpot Al Assistant (Ollama Integration)	Penpot (self- hosted)	On-device Al suggestions, screen generation, layout help	Fully local; privacy- preserving.
UX Pilot (Figma Plugin)	Figma	Prompt → mobile/web UI, wireframes, diagrams	Free plugin; generates pixel-perfect frames.
UXMagic (Figma Plugin)	Figma	Text → responsive UI, components, auto-layout, code export	Great for prototyping.
Galileo Al (Community Port)	Local (OSS)	Text → high-fidelity UI mockup (PNG)	Can be run locally; outputs design-ready images.
UI-TARS (Open- Source)	CLI/API	Prompt → React/Svelte component + Figma JSON	Bridges Al to code.
Macaw-LLM	Local	Multi-modal: generates images and text for UI elements	Requires setup.

Workflow for App Integration

- 1. Generate UI mockup using SDXL or Galileo Al.
- 2. Export PNG.
- 3. Import into Figma or Penpot.
- 4. Use UXMagic or UX Pilot to refine or extract components.

5. Export as React, Flutter, or Lottie animation.

4. Audio Generation – Sound Effects, Music, Voice

Generate audio clips, notification sounds, voiceovers, and music without cloud APIs.

Top Models

Model	Output	VRAM	Notes
Meta MusicGen	Music from text (10–30 sec clips)	4–8 GB	Trained on licensed music; ideal for background tracks.
MusicGPT (Rust wrapper)	Local GUI/CLI app for MusicGen	4+ GB	Precompiled binaries for Windows, Mac, Linux.
Stable Audio Open	Sound effects, samples (up to 47 sec)	6+ GB	From Stability AI; great for SFX.
AudioLDM 2	High-quality sound effects (44.1 kHz)	5+ GB	Prompt-based SFX generation.
Bark (suno/bark)	Text-to-speech with emotions, music, sound effects	4–6 GB	Supports 15+ languages; expressive voices.
Tortoise TTS	High-quality voice synthesis	6+ GB	Slow but realistic.
XTTS v2	Multilingual voice cloning (10 sec sample)	4+ GB	Popular for personalization.
Riffusion	Stable Diffusion fine-tuned on spectrograms → 5s music loops	3.8 GB	Visual music generation.
Coqui TTS	Open-source TTS with many language models	CPU- friendly	Offline, privacy-safe.
OpenAl Jukebox (MIT License)	Full song generation with style/artist conditioning	High VRAM	Research-grade; complex setup.

Workflow Example

python audioldm2.py "soft notification ding" \rightarrow ding.wav

Use in app as alert sound or integrate via LocalAI REST API.

5. Animation & Video Generation – GIFs, Lottie, CSS, 3D

Generate short animations for apps: loading spinners, transitions, micro-interactions, character motion.



Model/ Pipeline	Output	VRAM	Notes
AnimateDiff	PNG sequence → MP4/GIF/Lottie	8+ GB	Add motion to SD-generated images.
Stable Video Diffusion (SVD)	14–25 fps video from single image	10-12+ GB	Short clips (2–4 sec); ideal for app animations.
Genmo Mochi 1	Open-source text-to-video	8+ GB	Emerging alternative to Sora.
Wan 2.2	Text-centric animations (e.g., app demos)	6+ GB	Free local video generator.
LTX Video	Fast local video generation	8+ GB	Use with LTX Studio for app- ready clips.
MoMask	3D character animations (FBX/BVH)	4+ GB	From text or image prompts.
DeepMotion Animate 3D (CLI)	Rigged GLB → FBX animation	4+ GB	For 3D avatars in apps.
Dough	Precise Al animation steering	8+ GB	Open-source; GitHub install.
OpenToonz / Synfig	Traditional 2D animation with Al plugins	CPU/GPU	Free tools with Al-assisted frame generation.

Workflow Example

- 1. Generate static icon: sdxl.py "red bell icon" → bell.png
- 2. Animate: animatediff.py ——init bell.png ——prompt "bell shaking" → bell anim.mp4
- 3. Convert to Lottie using Bodymovin (After Effects) or LottieFiles converter.
- 4. Embed in app.
- **Description** LocalAI supports experimental video models via REST API.

Typical End-to-End Workflow for App Asset Generation

```
# 1. Generate PNG icon
python sdxl.py "flat style red notification bell icon, 512x512" > bell.png

# 2. Vectorize
python vectorize.py bell.png > bell.svg

# 3. Generate sound
python audioldm2.py "soft notification ding" > ding.wav

# 4. Animate
python animatediff.py --init bell.png --prompt "bell shaking gently" > bell_anim.mp4
```

5. Convert to Lottie (external tool)
lottie-convert bell_anim.mp4 bell.json

6. Import SVG and Lottie into Figma/Penpot → export for app

Hardware Recommendations

Tier	RAM	GPU	Use Case
Minimum	8 GB	GTX 1660 / RX 580 (4–6 GB VRAM)	Basic image/audio generation
Recommended	16 GB	RTX 3070 / 4060 Ti (8+ GB VRAM)	SDXL, AnimateDiff, MusicGen
Optimal	32 GB	RTX 4070 Ti / 4080 / 4090 (12–24 GB VRAM)	Video, 3D, high-res workflows

- **Figma/Penpot File Generation**: No model currently outputs native .fig or .penpot files. Workaround: generate PNG → import → manually recreate or use Al plugins.
- **Vector Generation**: Still in early stages. Most reliable path: raster → vectorize.
- Animation: Most models generate short clips (2–5 sec). Full video pipelines are experimental.
- Commercial Use: Most models are MIT, Apache-2.0, or CC-BY-NC. Always check individual licenses.
- Storage: Model files are large (2–12 GB each). Plan for 50–100 GB of free space.

🚀 Getting Started Checklist

- Choose a platform: ComfyUI, AUTOMATIC1111, or LocalAI.
- Install Python, PyTorch, CUDA (for NVIDIA).
- V Download a base model: SDXL 1.0 or Flux.
- Add extensions: ControlNet, AnimateDiff, IP-Adapter.
- Generate your first asset: "app login screen, modern UI, dark mode".
- V Post-process: vectorize, animate, export.
- Integrate into Figma, Penpot, or app code.

Future Outlook

- True Vector Diffusion Models expected to mature.
- Text-to-Figma/Penpot generators may emerge.
- Integrated Local Agent Platforms will unify text, image, audio, video, and code generation.
- 3D-to-Animation Pipelines will become more accessible.

This guide represents the state of open-source, locally-runnable generative AI as of August 2025, combining academic research, community projects, and industry tools into a unified, actionable resource.

Always check the specific license for each model for commercial use cases. This unified guide provides a clear path to building a complete, private, and cost-free generative Al asset pipeline on your local machine.