App: Local Image Generation (text2image_stablediffusion2)

- Requirements
- Installation
- Scaling
- App store
- Repository
- Known Limitations

The *text2image_stablediffusion2* app is one of the apps that provide image generation functionality in Nextcloud and act as an image generation backend for the Nextcloud Assistant app and other apps making use of the image generation functionality. The *text2image_stablediffusion2* app specifically runs only open source models and does so entirely on-premises. Nextcloud can provide customer support upon request, please talk to your account manager for the possibilities.

Requirements

- This app is built as an External App and thus depends on AppAPI v3.1.0 or higher
- Nextcloud AIO is supported
- We currently support NVIDIA GPUs and x86_64 CPUs
- CUDA >= v12.2 on your host system
- GPU Sizing

> ◦ A NVIDIA GPU with at least 8GB VRAM

- CPU Sizing

> - At least 8GB of system RAM
> - The more cores you have and the more powerful the CPU the better, we recommend 10-20 cores
> - The app will hog all cores by default, so it is usually better to run it on a separate machine

Installation

- Make sure the Nextcloud Assistant app is installed
- Install AppAPI and setup a Deploy Demon
- Install the "Local large language model" ExApp via the "Apps" page in the Nextcloud web admin user interface

Scaling

It is currently not possible to scale this app, we are working on this. Based on our calculations an instance has a rough capacity of 120 image requests per hour (each user request can be for multiple images). However, this number is based on theory and we do appreciate real-world feedback on this.

App store

You can also find the app in our app store, where you can write a review:
https://apps.nextcloud.com/apps/text2image_stablediffusion2

Repository

You can find the app's code repository on GitHub where you can report bugs and contribute fixes and features: https://github.com/nextcloud/text2image_stablediffusion2

Nextcloud customers should file bugs directly with our Support system.

Known Limitations

- The generated images are of a fixed resolution (512x512 pix), and the model does not achieve perfect photorealism
- The model cannot render legible text
- Faces and people in general may not be generated properly
- The results for certain image generation requests can be biased and may enforce stereotypes
- We currently only support languages that the underlying model supports; correctness of language use in languages other than English may be poor depending on the language's coverage in the model's training data
- Make sure to test the app for whether it meets the use-case's quality requirements
- Customer support is available upon request, however we can't solve false or problematic output, most performance issues, or other problems caused by the underlying model. Support is thus limited only to bugs directly caused by the implementation of the app (connectors, API, front-end, AppAPI)

□□□ #3
□ cocobook-sync □□□ 2026-10-28 16:28:52 UTC
□ cocobook-sync □□□ 2026-10-29 01:09:22 UTC