

Bringing Stable Diffusion to R: An R-First Implementation of Text-to-Image Generation

TL;DR

A native R implementation of Stable Diffusion SDXL, enabling text-to-image generation using TorchScript and the `{torch}` package.

Abstract

Stable Diffusion is a cutting-edge generative AI model for text-to-image synthesis, but its core implementations are limited to Python and rely heavily on object-oriented pipelines. This talk presents an R-first reimplementations of Stable Diffusion SDXL using the `{torch}` package and functional R design principles. Key components such as the UNet, VAE decoder, and text encoder were TorchScript-exported for ease of implementation, while the scheduling and sampling logic (e.g., DDIM, PNDM) is implemented directly in R.

The project aims to bring modern generative AI tooling to the R community with approachable syntax, functional composition, and extensibility for image-to-image, and eventually inpainting and text-to-video tasks. Attendees will learn how the package is organized, how to run inference locally with minimal setup, and how this approach could support reproducible and R-native generative workflows—without leaving the comfort of R.