

# Who watches the watchers?

LAION **describes itself** as a nonprofit organization with members worldwide, "aiming to make large-scale machine learning models, datasets and related code available to the general public." Its data can be used in various projects, from facial recognition to computer vision to image synthesis.

For example, after an AI training process, some of the images in the LAION data set become the basis of Stable Diffusion's **amazing ability** to generate images from text descriptions. Since LAION is a **set of URLs** pointing to images on the web, LAION does not host the images themselves. Instead, LAION **says** that researchers must download the images from various locations when they want to use them in a project.



**Enlarge** / The LAION data set is replete with potentially sensitive images collected from the Internet, such as these, which are now being integrated into commercial machine learning products. Black bars have been added by Ars for privacy purposes.