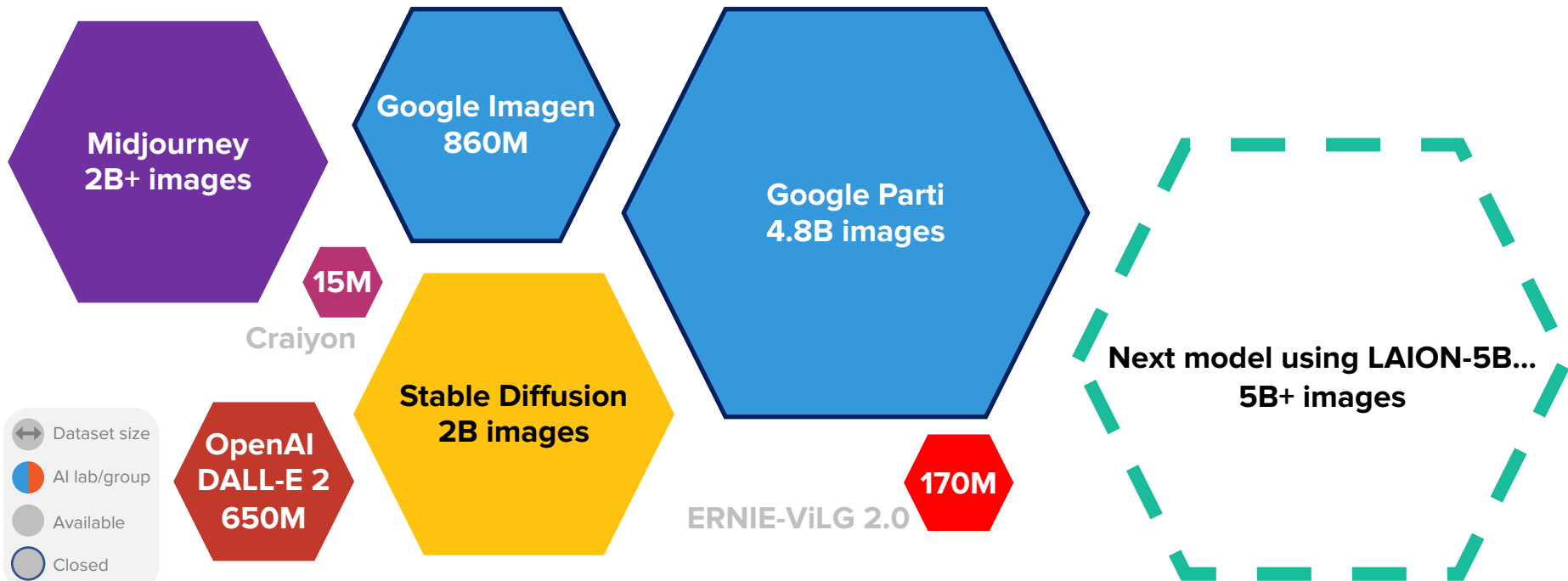


TEXT-TO-IMAGE TRAINING DATASET SIZES. 2022–



Sources: MidJourney: 2B+ images "trained over billions of images." https://www.theregister.com/2022/08/01/david_holz_midjourney/, Dall-E 2: 650M images "we sample from the CLIP and DALL-E datasets (approximately 650M images in total)" <https://cdn.openai.com/papers/dall-e-2.pdf> Craiyon: 15M images <https://arxiv.org/abs/2208.09333> Google Imagen: 860M images "We train on a combination of internal datasets, with = 460M image-text pairs, and the publicly available Laion dataset, with = 400M image-text pairs." <https://arxiv.org/abs/2205.11482> SD: 2B images "237k steps at resolution 256x256 on laion2B-en" <https://github.com/CompVis/stable-diffusion/>, Google Parti: 5B images "The data includes the publicly available LAION-400M dataset; FIT-400M, a filtered subset of the full 1.8 billion examples used to train the ALIGN model; JFT-4B dataset, which has images with text annotation labels" <https://arxiv.org/abs/2206.10782> Beeswarm/bubble plot, sizes mostly to scale. Selected highlights only. Alan D. Thompson, September 2022. <https://life architect.ai/>

