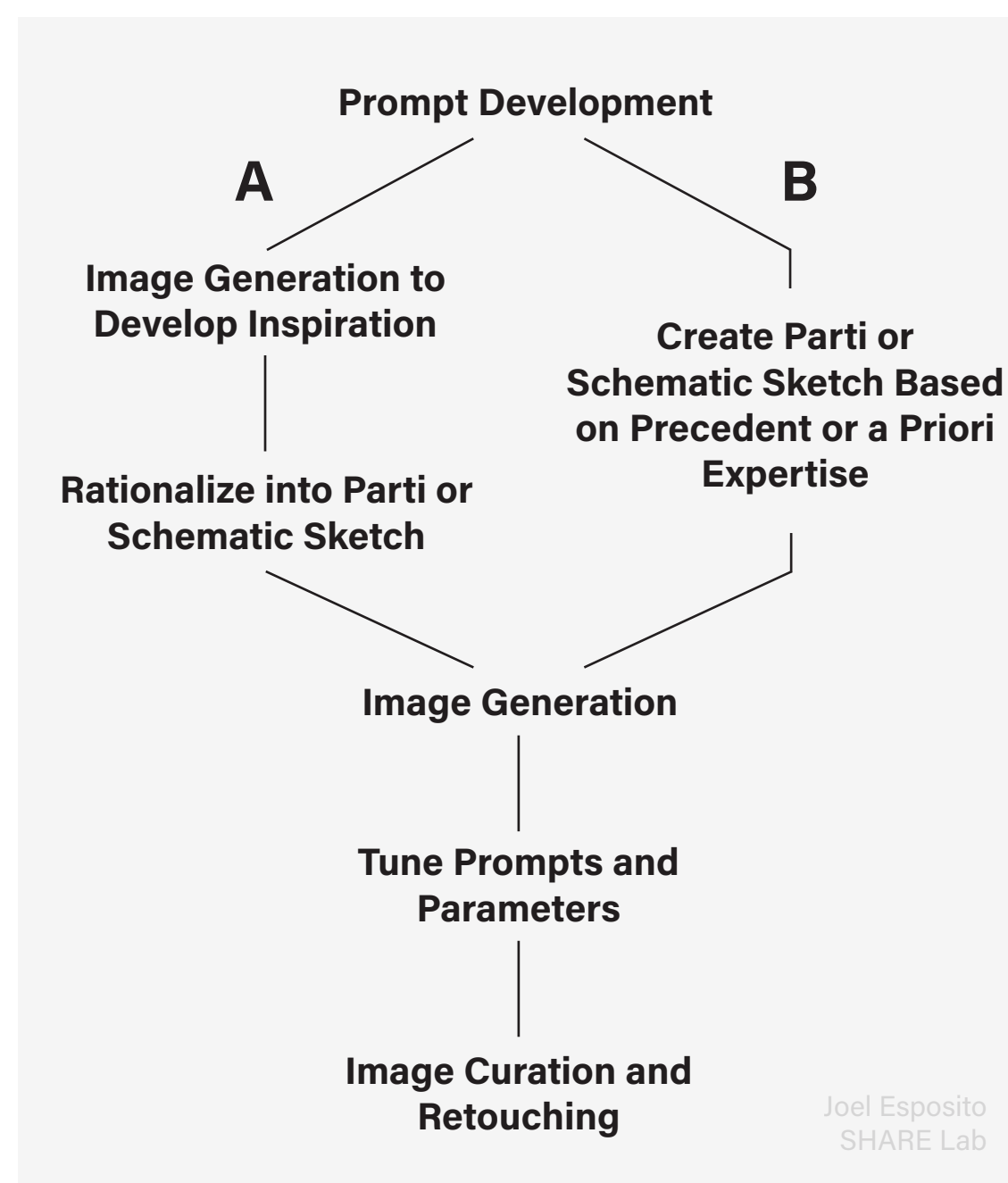


Controlling ControlNet

Joel Esposito (JaxLab, 2023)

Using a Latent Diffusion Model to Augment the Schematic Design Process

Workflow

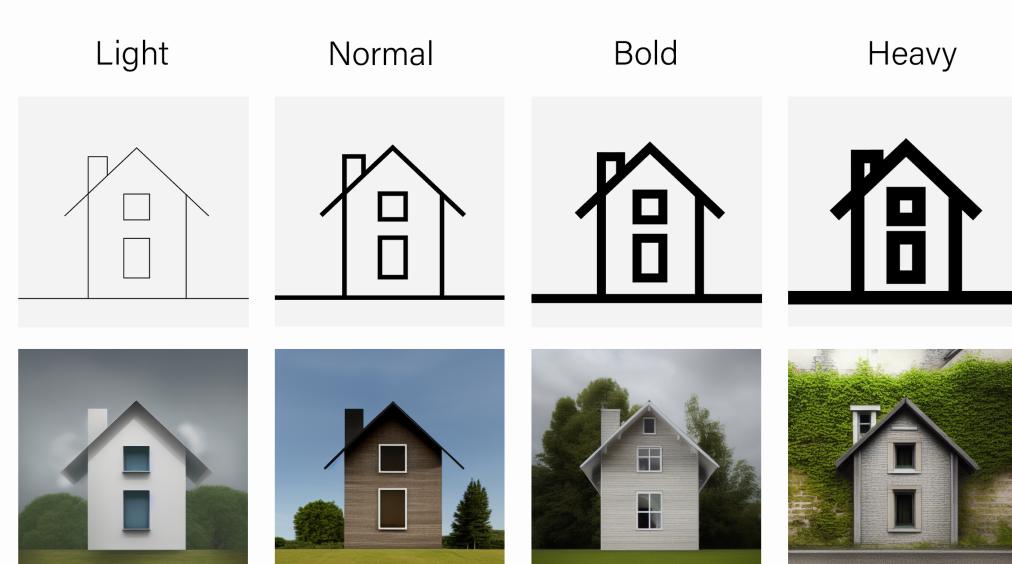


An experiment was designed to compare varying combinations of image LOD and prompt LOD. Note how the level of inference achieved by the diffusion model becomes increasingly sophisticated, albeit in different domains, as each axis approaches a high LOD. For example, contextually-appropriate shading appears with the highest prompt specificity, while fine details like fenestration are more accurately rendered with a high detail input sketch.

Parti Matrix Test

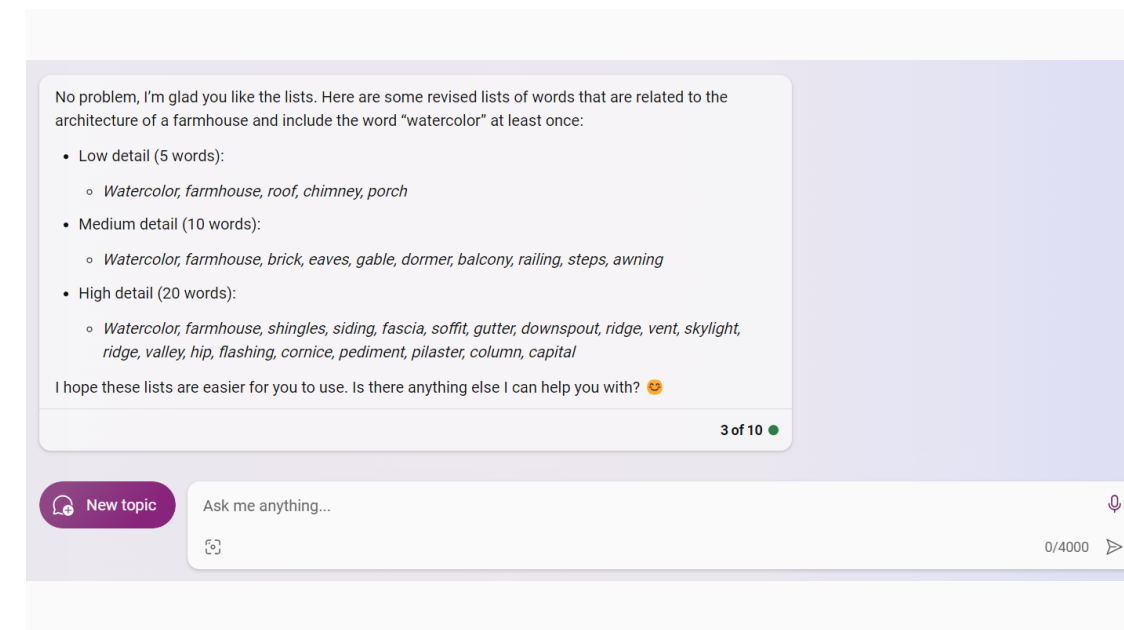


Stroke Weight Testing



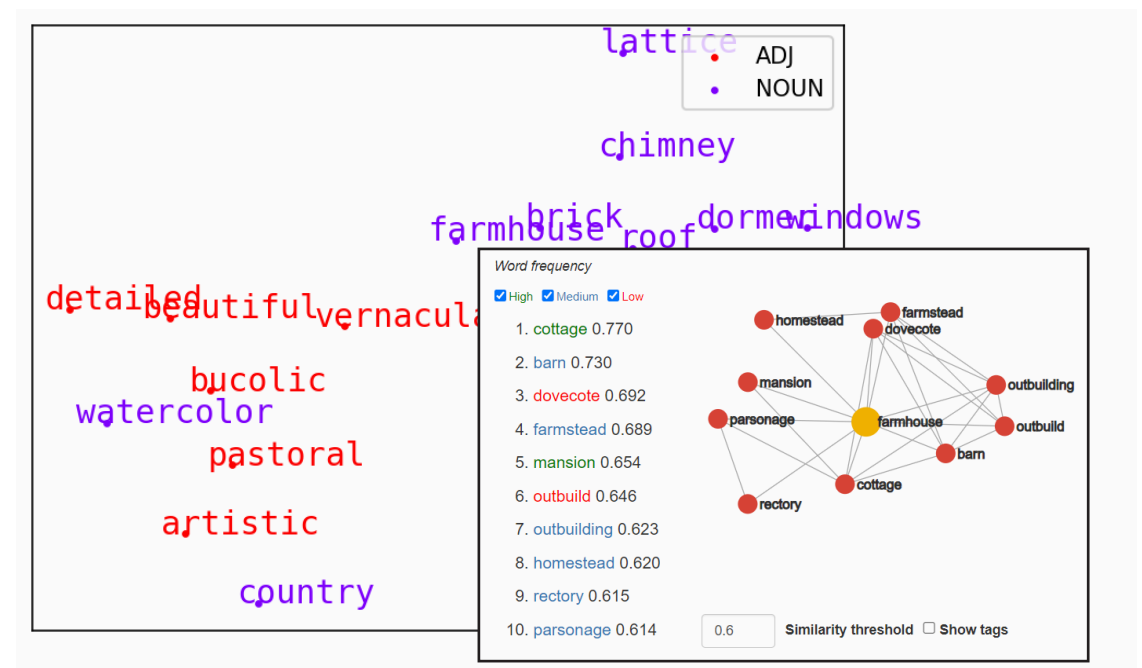
The correlation between stroke weight and level of detail is noteworthy. Further testing will be required to determine the exact nature of how stroke weight interfaces with other input variables.

Chatbot Integration



Using Bing AI Chat (a large English language model based chatbot) to rapidly develop robust lists of architectural vocabulary for further experimentation.

Semantic Analysis



Future experiments can include leveraging other variables such as the semantic proximity of words within a given prompt.

Workflow "A" (Studio Work)



Example of workflow "A" used to rapidly develop spatial ideas, building massing, and context. Resultant sketches can be brought back into the diffusion model for rendering or reused in ControlNet. The feedback loop between designer and the diffusion model augments a priori expertise of the designer with speed, fidelity, and breadth offered by the image generator.



Copyright Considerations

Artifacts from vast image data sets sometimes include agglomerated artist signatures. This raises important questions about the "opt-out only" nature of content that is utilized by image generators and the essential role that the data set plays in AI systems. The discourse surrounding copyright and image generation is rapidly evolving. At the present time, AI images are a catch-22 from a business perspective—they enable much faster workflows, but the raw output is not eligible for copyright. Direct manipulation of the output by the artist and/or more direct involvement of the artist during image creation (e.g. NVIDIA Canvas) will almost certainly be required to establish copyright over the image.



Looking Forward

The English language version of XKool AI Cloud is one example of a recently launched platform that offers a user-friendly interface as well as a robust feature set that can facilitate the aforementioned workflows with a lower bar to entry (faster learning curve).