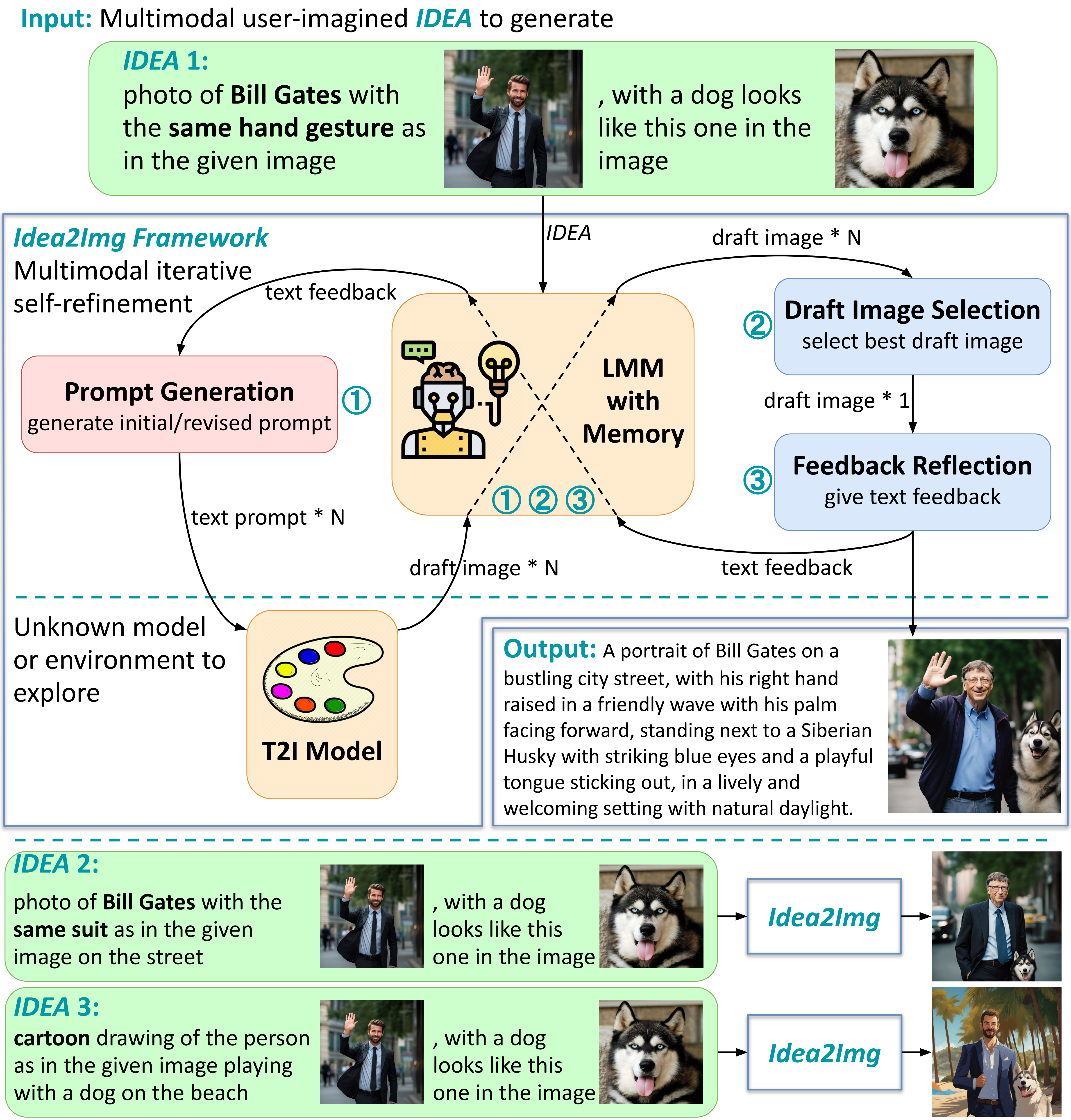

Input: Multimodal user-imagined *IDEA* to generate

IDEA 1: photo of **Bill Gates** with the **same hand gesture** as in the given image , with a dog looks like this one in the image

Idea2Img Framework

Multimodal iterative self-refinement

Output: A portrait of Bill Gates on a bustling city street, with his right hand raised in a friendly wave with his palm facing forward, standing next to a Siberian Husky with striking blue eyes and a playful tongue sticking out, in a lively and welcoming setting with natural daylight.

IDEA 2: photo of **Bill Gates** with the **same suit** as in the given image on the street , with a dog looks like this one in the image

IDEA 3: **cartoon** drawing of the person as in the given image playing with a dog on the beach , with a dog looks like this one in the image