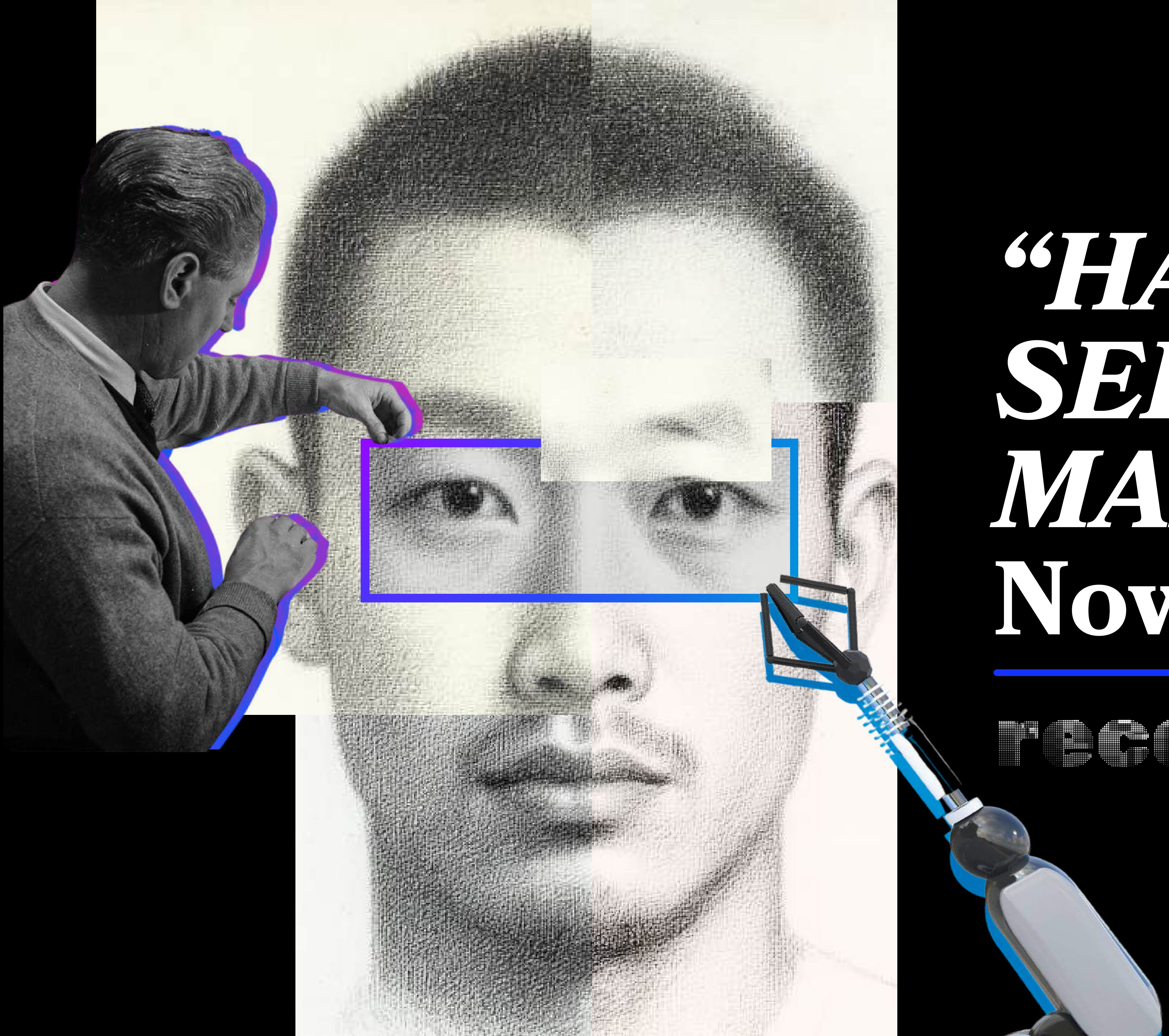


“HAVE YOU SEEN THIS MAN?” Now, AI has.

recogn-eyes

Facial Composites are Crucial and Effective

Composites provide visual leads, prompting urgent public assistance when necessary.

67%

of sketch composites were recognized by someone.

76%

of the solved cases had the offender identified through a lead from the sketch.

41%

of cases are eventually cleared with an arrest.

Composites aren't just for narrowing down potential suspects based on descriptions. Sometimes, no surveillance footage exists.

Sources:

1. Citation:International Association for Identification. (n.d.). JFI Abstracts from 2018–2021. Retrieved April 17, 2025, from https://www.theiai.org/docs/JFI_Abstracts_from_2018-2021.pdf#:~:text=artists%20completed%20the%20survey%20and,of%20offenders%20were%20identified%20through

... But current processes are resource-intensive

But current template-assembly processes are slow and resource-intensive

IO interviews witness for rapport, cognitive recall, and description of features.

→ Feature-by-feature composite is assembled using **preset templates**

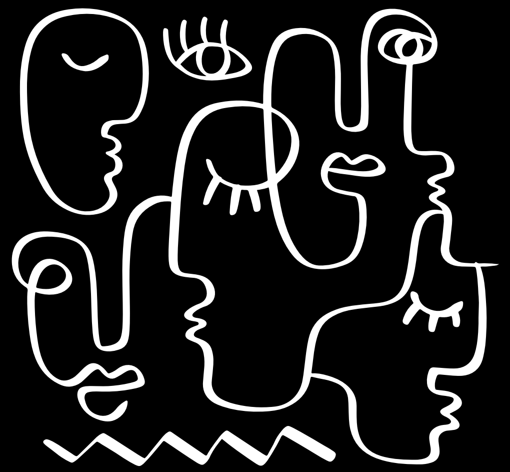
→ IO makes **iterative changes manually**, with feedback from witness.

2-3 hours per composite, not factoring time taken for the interview process.
The longer it takes, witness memory loss reduces accuracy.

Sources:
1. "Whole face" software to improve composite sketches of criminals | Homeland Security Newswire. (n.d.). https://www.homelandsecuritynewswire.com/whole-face-software-improve-composite-sketches-criminals
2. Chandler, N. (n.d.). How police sketches work. HowStuffWorks. Retrieved April 17, 2025, from https://people.howstuffworks.com/police-sketch.htm

Findings on Composite Accuracy

Composites provide visual leads, prompting urgent public assistance when necessary.



Feature-by-feature method used to construct composites from templates is at odds with the natural, global process of face recognition

Hyperrealism can create false familiarity, causing witnesses to mistake composites for actual memories, especially if the image looks too plausible

Facial composite need not be photorealistic, but must be familiar enough for recognition and lead generation

Sources:

1. Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. The Quarterly Journal of Experimental Psychology, 46(2), 225–245. https://doi.org/10.1080/14640749308401045JOV+2Wikipedia+2Sp
2. Zajac, R., & Karageorge, A. (2009). The negative effect of verbal overshadowing in face composite construction: A research note. Applied Cognitive Psychology, 23(1), 84–91. https://doi.org/10.1002/acp.1441
3. Frowd, C. D., Bruce, V., & Hancock, P. J. B. (2008). Improving the quality of facial composites using a holistic cognitive interview. Journal of Experimental Psychology: Applied, 14(3), 276–287. https://doi.org/10.1037/1076-898X.14.3.276

Common Limitations of Computer Composites

Computer generated composites often meet skepticism due to...

Unclear Workflow Integration

US Homeland Security Taxonomy Report:

*“Tools need to **fit into** training, systems engineering, and decision support workflows.”*

- Workflows that **neglect interviews** by **cutting out investigative officers** result in **lower accuracy**
- AI needs to fit *within* forensic workflows

Early Technology Constraints

Legacy police composite software relies heavily on feature-based recall.

- **Whole-face recognition** (not feature-by-feature recall) **yields better results** by helping users choose what "feels familiar."

Concerns with AI impact

Bias concerns due to **small training datasets** used for AI models in the past.

Low **Transparency** of public online models.

Privacy and **Security** concerns with online tools.

Sources:

1. Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. (2018). High-resolution image synthesis and semantic manipulation with conditional GANs. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 8798–8807.

Analysis of Existing Market Solutions

	Rendering Method	Privacy-Ready	Bias/Representation	Cognitive Fit	Designed for Police
Our Solution	**Automated, 2D rendering**	✓	**Mitigated by Multi-model + LoRA**	**Still involves recall-process, but relies on recognition***	✓
Composite Software					
KLIM 3D	3D Modelling	✓	Mitigated	Relies on feature recall	✓
FACES (Legacy)	Manual, 2D template assembly	✓	Mitigated, Restricted to available templates	Relies on feature recall	✓
FACES (IQ Biometrix)	Manual, 2D template assembly	✓	Mitigated, Restricted to available templates	Relies on feature recall	✓
SketchCop	Manual, 2D template assembly	✓	Mitigated	Still involves recall-process, but relies on recognition	✓
Image Generators					
GPT-4o (Vision)	Automated, 2D rendering	✗	Limited control	NA	✗
Midjourney / DALL·E	Automated, 2D rendering	✗	Known issues	NA	✗

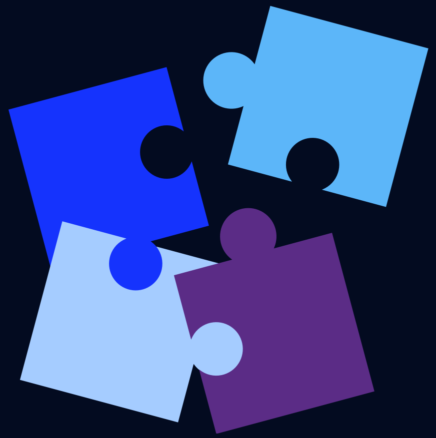
RecognEyes overcomes these challenges

	Rendering Method	Privacy-Ready	Bias/Representation	Cognitive Fit	Designed for Police
Our Solution	**Automated, 2D rendering**	✓	**Mitigated by Multi-model + LoRA**	**Still involves recall-process, but relies on recognition***	✓
Composite Software					
KLIM 3D	3D Modelling	✓	Mitigated	Relies on feature recall	✓
FACES (Legacy)	Manual, 2D template assembly	✓	Mitigated, Restricted to available templates	Relies on feature recall	✓
FACES (IQ Biometrix)	Manual, 2D template assembly	✓	Mitigated, Restricted to available templates	Relies on feature recall	✓
SketchCop	Manual, 2D template assembly	✓	Mitigated	Still involves recall-process, but relies on recognition	✓
Image Generators					
GPT-4o (Vision)	Automated, 2D rendering	✗	Limited control	NA	✗
Midjourney / DALL·E	Automated, 2D rendering	✗	Known issues	NA	✗

First **automated rendering** among composite softwares. Mitigates **bias and representatives** unlike other image generation tools, and **tailored towards investigation use cases**.

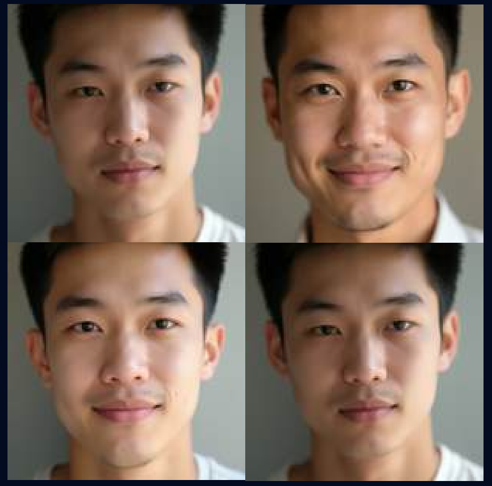
RecognEyes is built with real-world usability in mind

In the public sector, trust is paramount. We understand there is no room for failure.



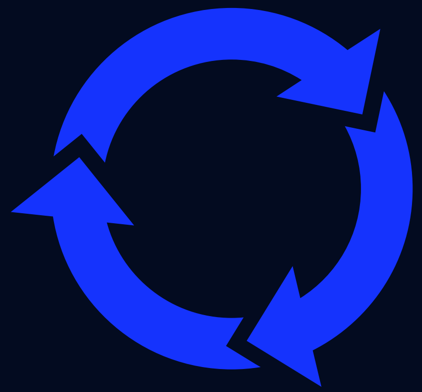
Plug and Play

- Built with HITL (human-in-the-loop) workflows in mind
- Swap models and LoRAs based on profiling needs
- Supports diverse combinations using open-source tools
- FLUX used as base model due to strong portrait performance



4 Choose 1

- Research shows it's more effective to identify from multiple similar faces than rely on one high-accuracy image
- Multiple outputs are shown to the witness for best match
- Achieved by varying model and LoRA combinations, plus pipeline configuration settings

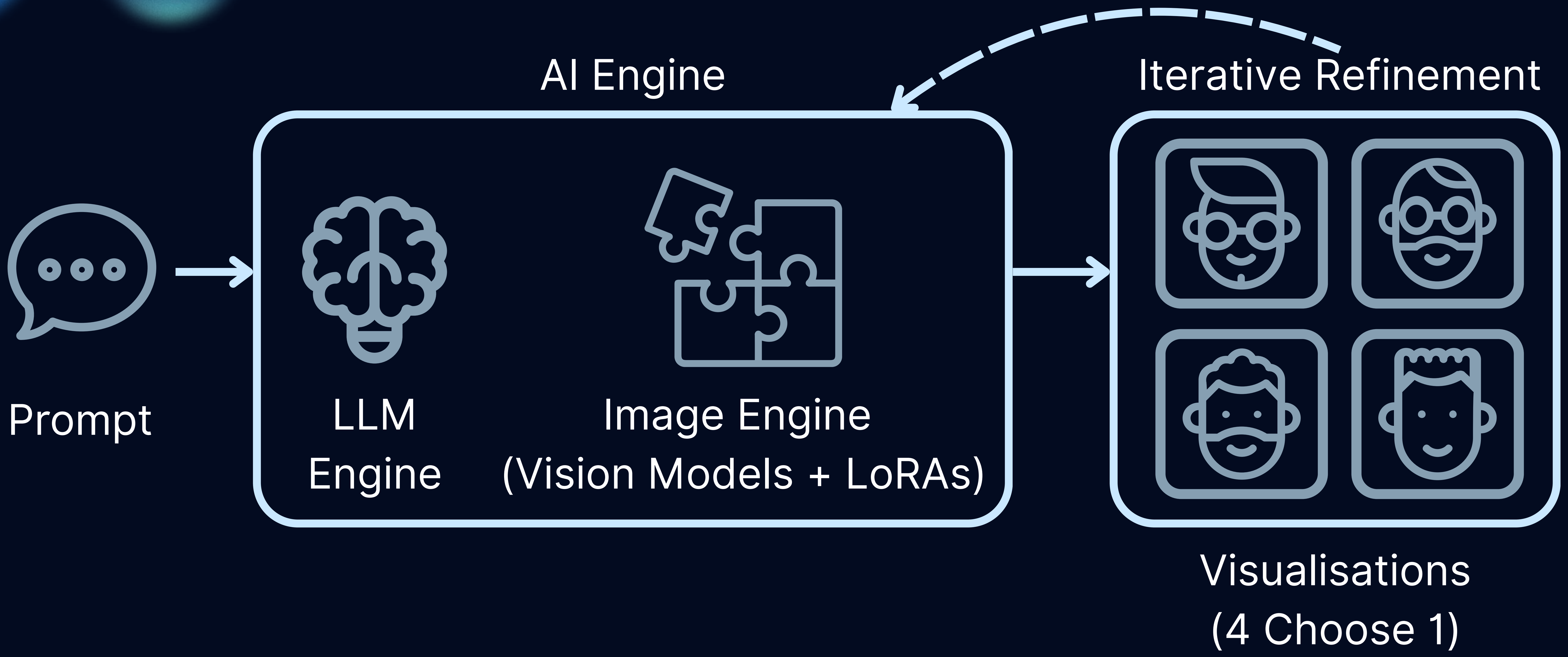


Refine and Repeat

- Witness then selects the closest match
- The selected image becomes the base for further refinement
- Additional details can then be added iteratively
- Refinement uses an LLM layer that rephrases inputs with semantic and synonym variation (inspired by the concept of some linguistic descriptions being subjective)

Sources:
1. Liu, Y., Zhang, K., Li, Y., Yan, Z., Gao, C., Chen, R., ... & Sun, L. (2024). Sora: A review on background, technology, limitations, and opportunities of large vision models. arXiv preprint arXiv:2402.17177.
2. Zheng, C., Lan, Y., & Wang, Y. (2025). Lanpaint: Training-Free Diffusion Inpainting with Exact and Fast Conditional Inference. arXiv preprint arXiv:2502.03491.
3. Dowsett, A. J., & Burton, A. M. (2015). Face learning with multiple images leads to fast acquisition of familiarity for specific individuals. Quarterly Journal of Experimental Psychology, 68(9), 1715–1726. https://doi.org/10.1080/17470218.2014.1003949

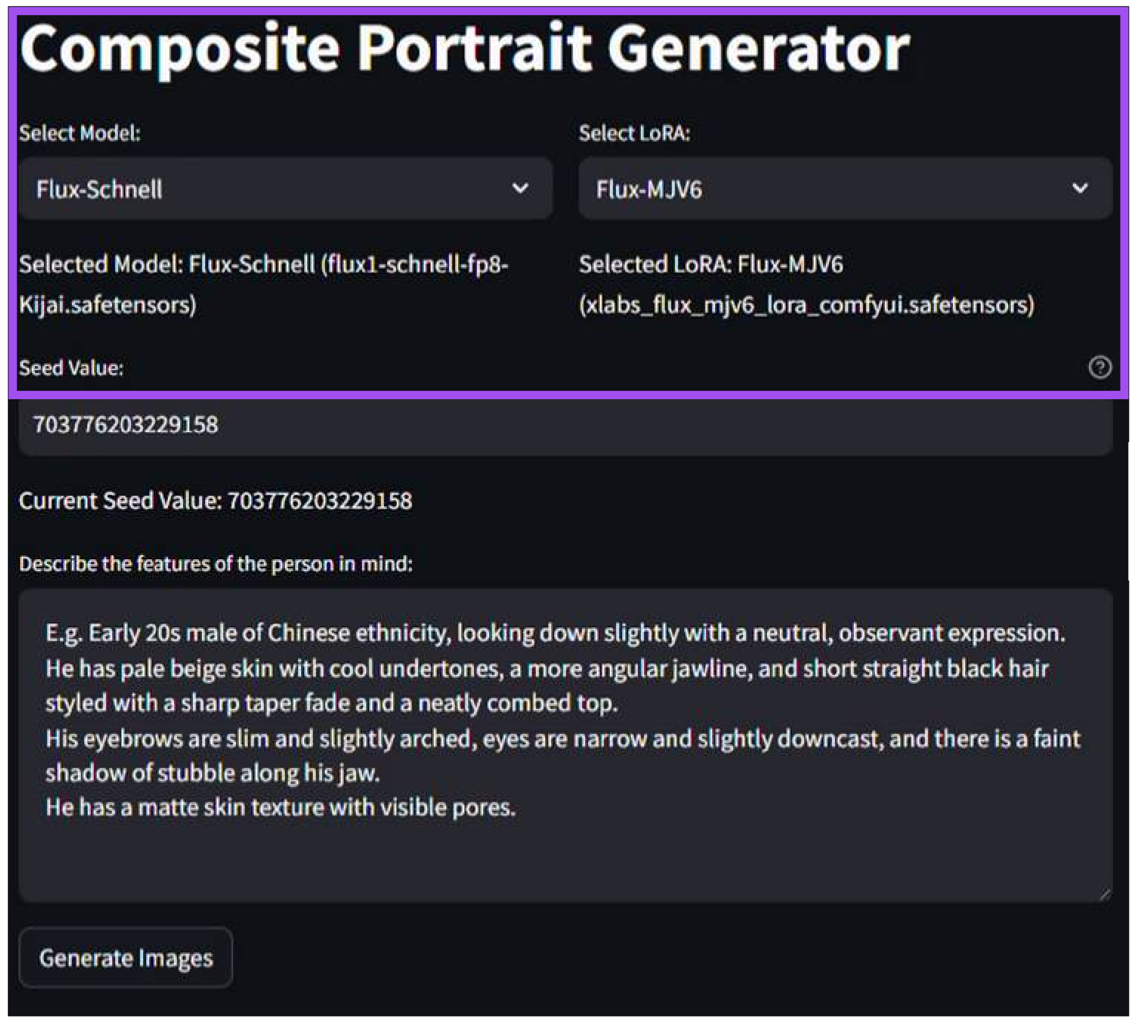
Architecture

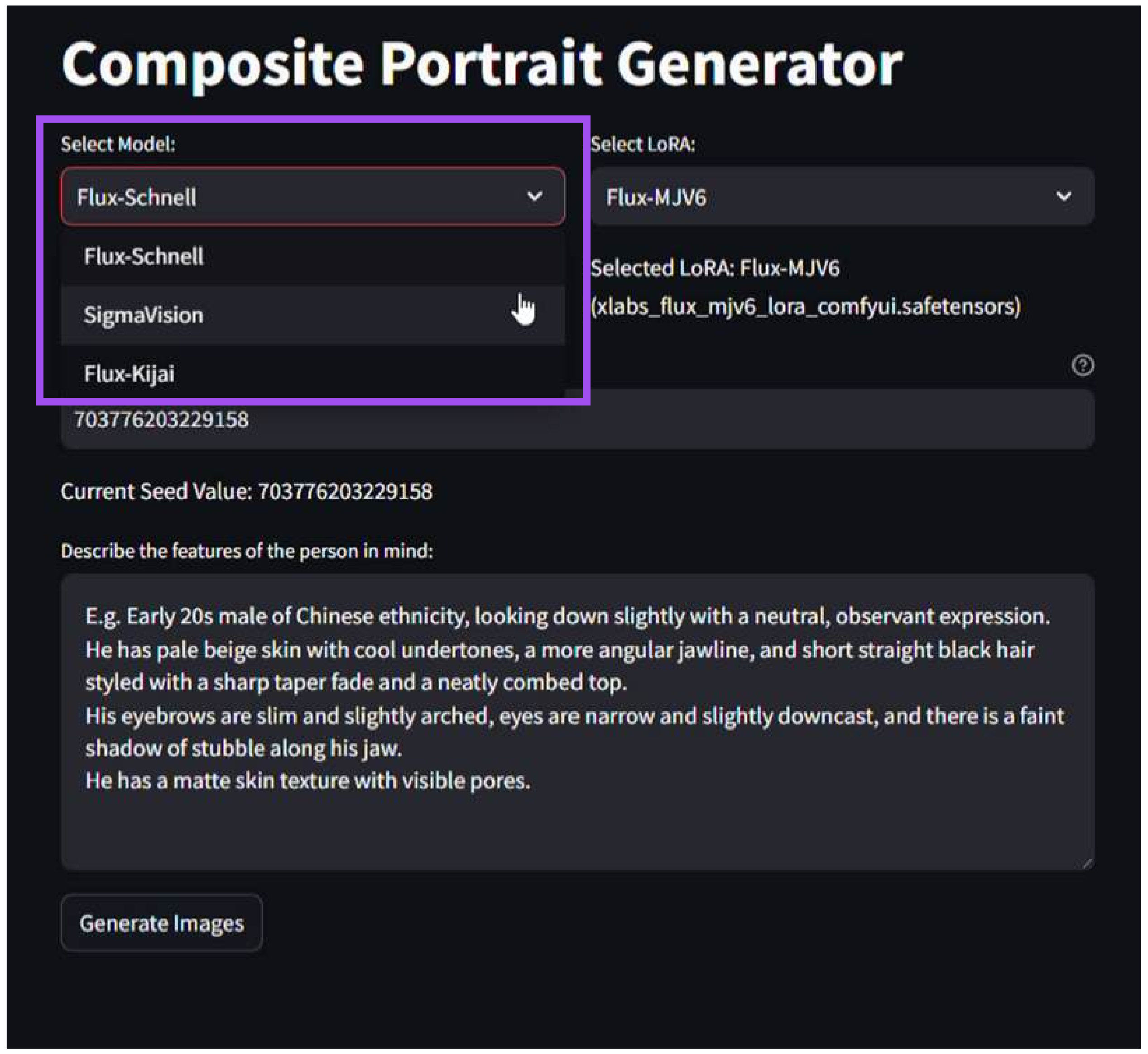


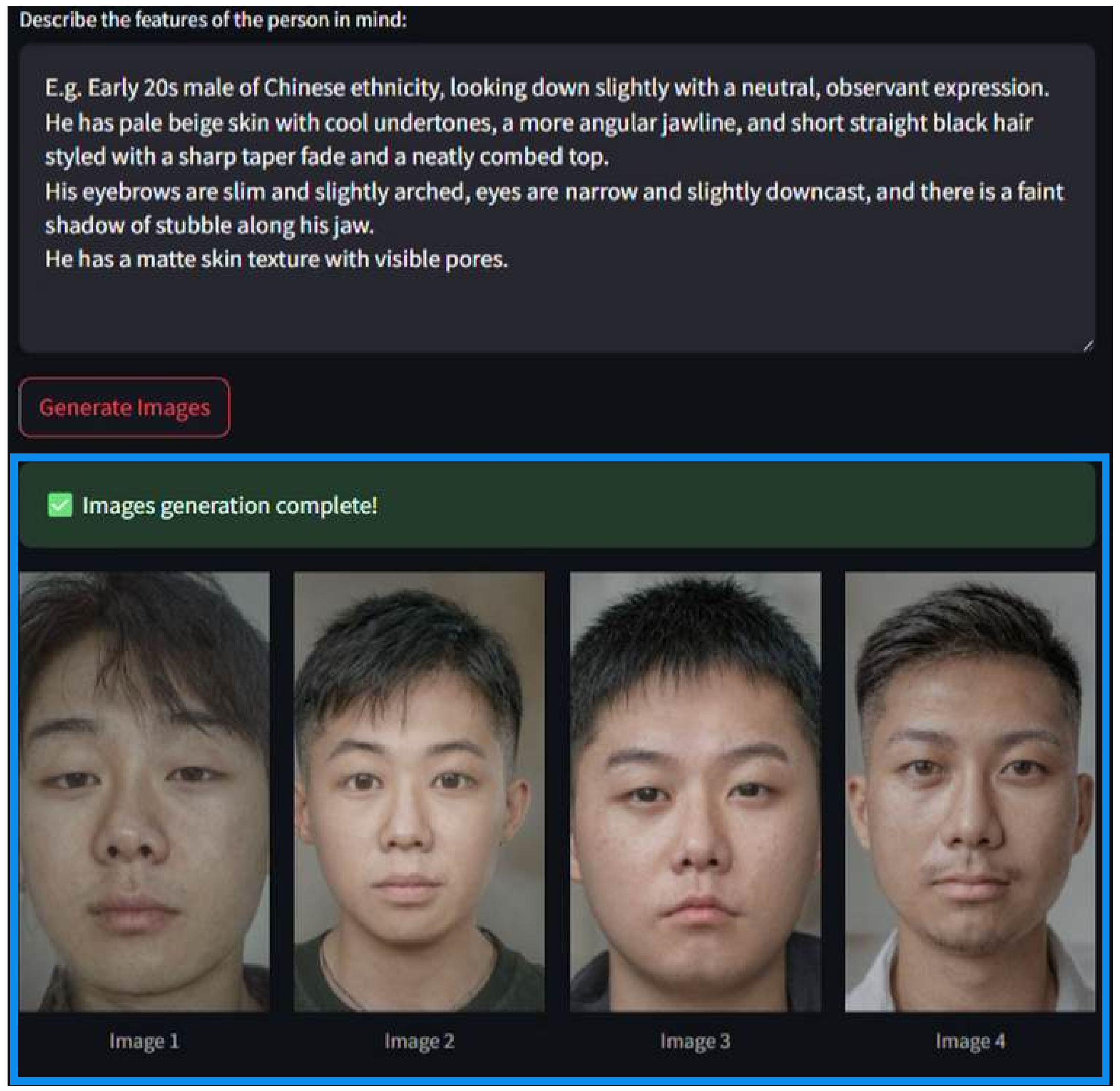
Sources:
1. Tero Karras, Samuli Laine, and Timo Aila. (2019). A style-based generator architecture for generative adversarial networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 4401–4410.

Overview (Part 1)

1) Plug and Play Models
2) 4 Choose 1 (Produced Visualisations)
3) Iterative Refinement



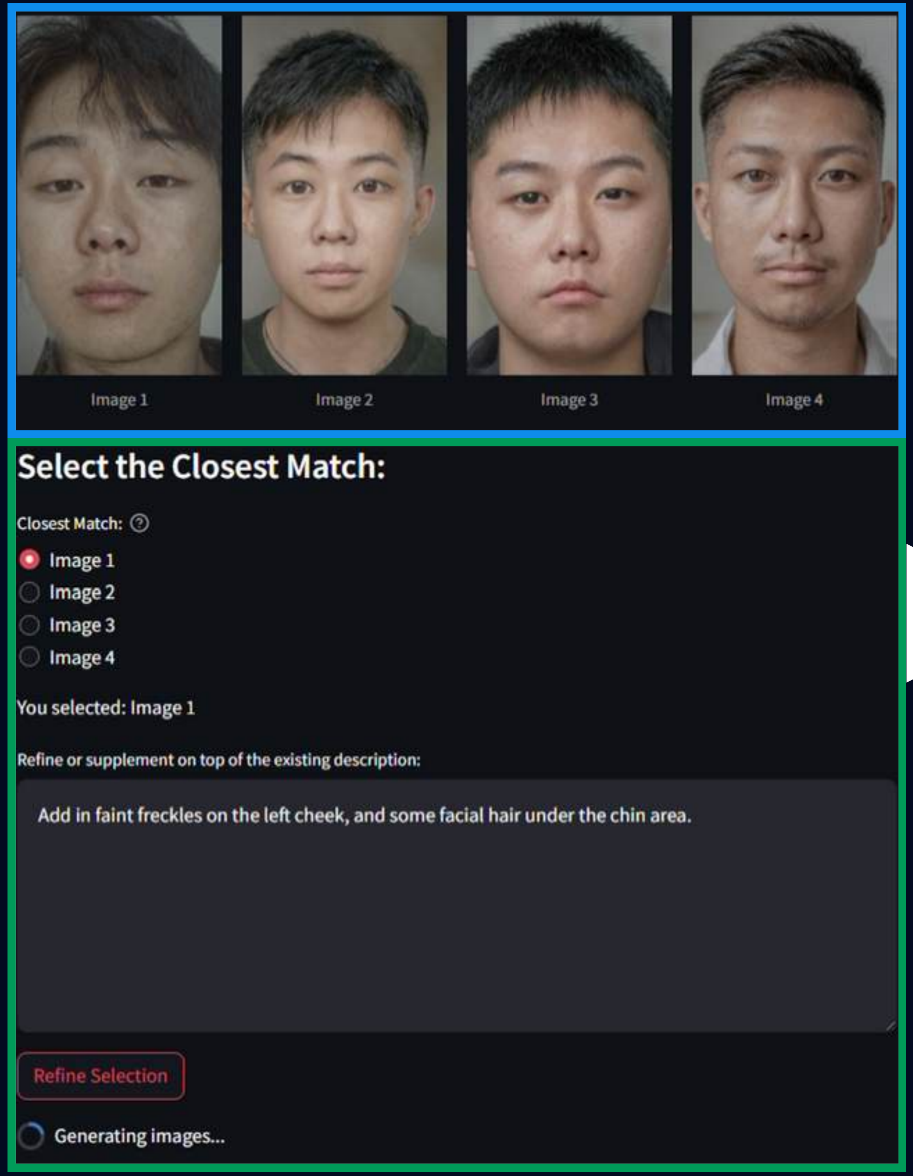


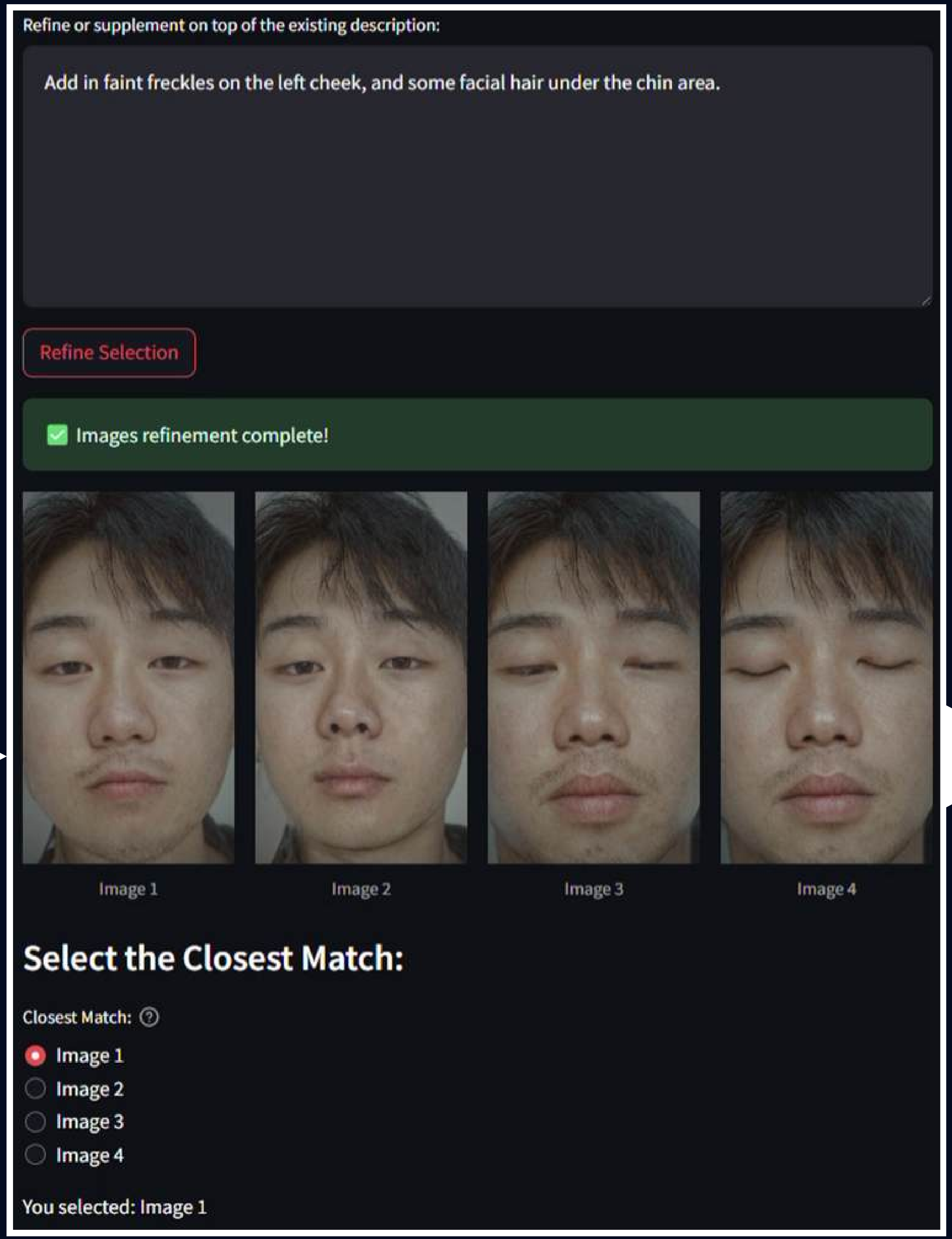


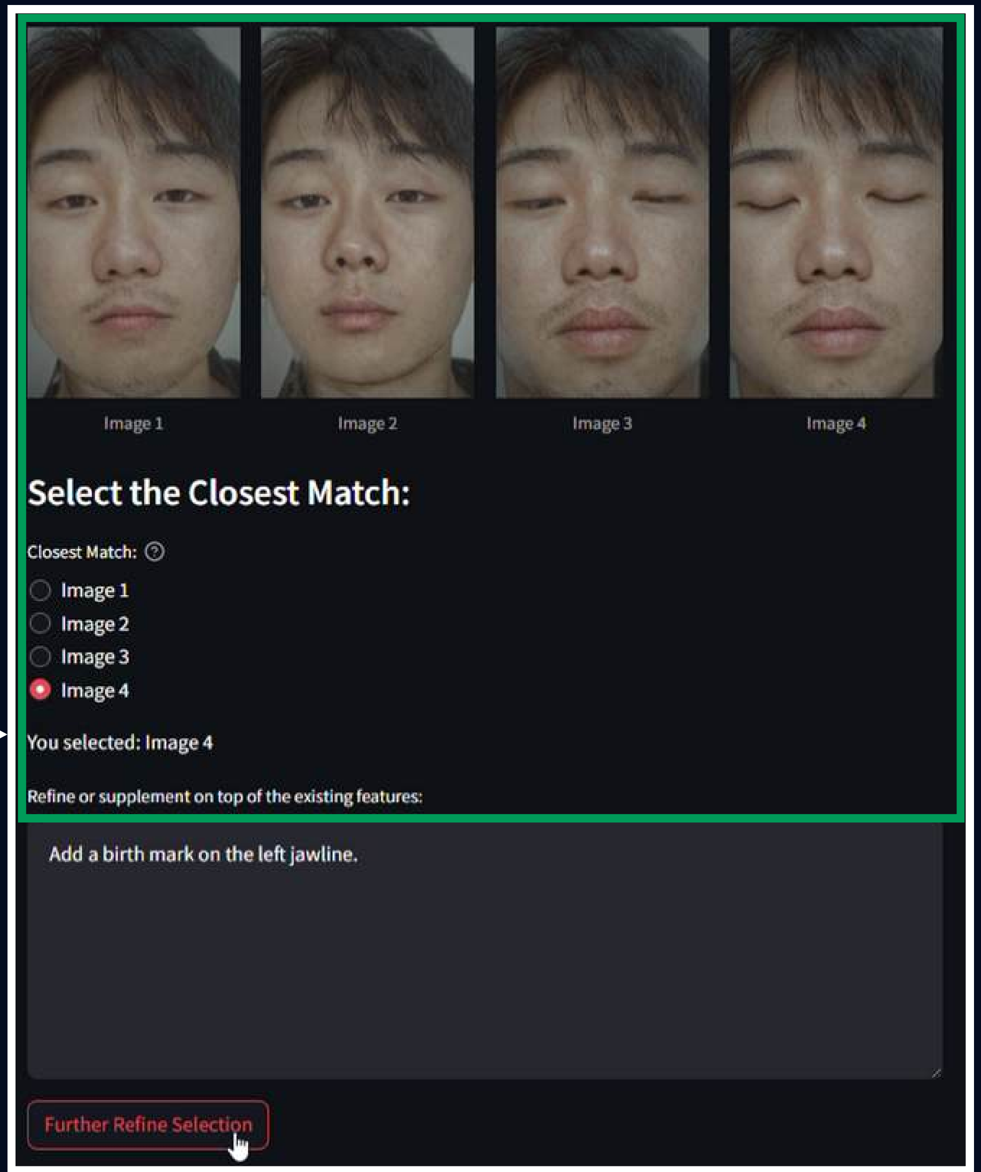
Watch our demo: https://youtu.be/4fHYJuApNxM

Overview (Part 2)

1) Plug and Play Models
2) 4 Choose 1 (Produced Visualisations)
3) Iterative Refinement







Watch our demo: https://youtu.be/4fHYJuApNxM

Production Considerations

Tech for public good - GenAI techniques worth exploring when scaling to the masses.

1 Quantization Techniques

Reduce precision of model weights and shrink image resolutions. This significantly decreases* the model's size and memory footprint, leading to faster inference speeds and enabling deployment on devices with limited resources or in secure offline environments.

**Roughly inversely proportional relationship, e.g., image of 4K to 1080P quality would roughly equate to 4x benefits in memory savings and inference speeds.*

2 Knowledge Distillation

Also referred as Teacher-Student concept, train a smaller model from a larger teacher model using prompts and outputs from high-quality models to fine-tune lighter architectures. This allows for relatively superior performance to be achieved with smaller models that are significantly more cost-effective, faster for real-time applications, and easier to secure.

3 Matryoshka Representation Learning

Instead of the whole image, represent generated images using compressed embeddings comprised of various layers which offer different levels of detail. This significantly reduces storage requirements and enables fast similarity searches based on visual features, making it highly valuable for internal indexing, case tracking, and matching descriptions.

4 LoRA Adapters

Utilise small, lightweight modules that can be plugged into base models to customize based on the given scenario needs. This approach keeps the core system lean and modular, allowing for adaptability to diverse requirements and demographic profiles without the need to store large, specialized models.

Sources:

1. Gholami, A., Kim, S., Yao, Z., Mahoney, M. W., & Keutzer, K. (2021). A survey of quantization methods for efficient neural network inference. arXiv. https://arxiv.org/abs/2103.13630
2. Zhou, Y., Zhang, J., & Yang, Y. (2023). Teacher-student architecture for knowledge distillation: A survey. arXiv. https://arxiv.org/abs/2308.04268
3. Beaumont, R., Bansal, A., & Goyal, P. (2022). Matryoshka representation learning. arXiv. https://arxiv.org/abs/2205.13147

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, L., ... & Chen, W. (2021). LoRA: Low-rank adaptation of large language models. arXiv. https://arxiv.org/abs/2106.09685

Design Considerations

Balancing trust, interpretability, and practical outcomes in this high-stake AI application.

Designed for HITL* - "Guiding Assistant, Not A Replacement"

**HITL = Human-In-The-Loop*

Many AI solutions are built for full autonomy and leave the human out of the process.

However, for investigative applications, this approach is risky. The black box nature of generative AI makes it difficult to trace or justify how outputs are produced.

In high-criticality scenarios, human-in-the-loop workflows ensure accountability, context, and trust.

Rethinking Accuracy as a Success Metric

Accuracy is a common benchmark for evaluating AI systems. However, it is not the most appropriate measure for our intended use case.

Since hyperrealistic outputs can interfere with memory recall, our approach prioritises enabling witnesses to identify the most familiar or closely resembling face from a set of options.

Accordingly, we evaluate success based on time to completion rather than visual accuracy.

Sources:
1. Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Horvitz, E. (2022). Human-in-the-loop machine learning: A state of the art. Artificial Intelligence Review, 55(2), 1-45. https://link.springer.com/article/10.1007/s10462-022-10246-w
2. Kumar, S., & Singh, A. (2024). Rethinking reward model evaluation: Are we barking up the wrong tree? arXiv. https://arxiv.org/abs/2410.05584arXiv

Implementation Roadmap

Ready from Day One, Here for the Long Run.

Legend
- Version 0 for immediate testing
- Version 0.2 for in-house integration
- Version 1 and beyond for iterative testing and refinement

Phase	Phase 1	Phase 2	Phase 3
Product	**Pilot testing** With selected locations Initial A/B testing Gather preliminary feedback and comments	**Further testing (In stages)** Expand locations Engage with potential partner collaborators (e.g., ICA) Commence data collection	
Implementation	**Production capabilities** Compatible with local intranets and dev environments Confirm Cloud (GCC) or on-premise deployment Adjust GenAI models to suit offline capabilities	**CI/CD** Model Refinement and Finetuning Model refresh based on existing open source solutions Identify better performing Model, LORA pairs and offer as recommendations	
Research	**Considerations** NLP semantic and synonym refinement Model parameters experimentation and finetuning Exploring alternative image models and LORAs		

Meet The Team!

Tech enthusiasts passionate in harnessing AI for public good.

Jared
NUS MSc. Business Analytics
ex-IDEMIA, GovTech
Project/AI Lead

Malcolm
SIM-UOL Digital Innovation
ex-GovTech
AI Engineer

Chloe
NUS Business Administration
KPMG
Product Manager

Sibikrishnan
SIM-UOL Computer Science
EY
AI Engineer

Jolin
NUS Data Science and Analytics
GovTech
Research Scientist

Xin Tian
NUS Data Science and Economics
GovTech
Data Scientist

Advisors

Gene Wuu
SVP & CTO of UTStarcom

Sachin Tonk
Deputy CDO of GovTech

References

Black Forest Labs. (2024). FLUX. Retrieved from
https://github.com/black-forest-labs/flux

Chen, S. Y., Su, W., Gao, L., Xia, S., & Fu, H. (2020). Deep generation of face images from sketches. arXiv preprint arXiv:2006.01047.
Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 1125–1134.

Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 4401-4410).

Liu, Y., Zhang, K., Li, Y., Yan, Z., Gao, C., Chen, R., ... & Sun, L. (2024). Sora: A review on background, technology, limitations, and opportunities of large vision models. arXiv preprint arXiv:2402.17177.

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. (2017). Image-to-image translation with conditional adversarial networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 1125–1134.

Patsorn Sangkloy, Jingwan Lu, Chen Fang, Fisher Yu, and James Hays. (2017). Scribbler: Controlling deep image synthesis with sketch and color. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 5400–5409.

Sharma, A., Yu, A., Razavi, A., Toor, A., Pierson, A., Gupta, A., ... Chen, Y. (2024). Veo. Retrieved from https://deepmind.google/technologies/veo/

Tero Karras, Samuli Laine, and Timo Aila. (2019). A style-based generator architecture for generative adversarial networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 4401–4410.

Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. (2018). High-resolution image synthesis and semantic manipulation with conditional GANs. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 8798–8807.

McLean, C. R., Jain, S., & Lee, Y. T. (2008). A Taxonomy of Homeland Security Modeling, Simulation, and Analysis Applications | NIST. National Institute of Standards and Technology Gaithersburg. https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=824440

Miao Wang, Guo-Ye Yang, Ruilong Li, Run-Ze Liang, Song-Hai Zhang, Peter M Hall, and Shi-Min Hu. (2019). Example-guided style-consistent image synthesis from semantic labeling. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 1495–1504.

“Whole face” software to improve composite sketches of criminals | Homeland Security Newswire. (n.d.). https://www.homelandsecuritynewswire.com/whole-face-software-improve-composite-sketches-criminals

Zheng, C., Lan, Y., & Wang, Y. (2025). Lanpaint: Training-Free Diffusion Inpainting with Exact and Fast Conditional Inference. arXiv preprint arXiv:2502.03491.

Thank You