

[멋쟁이 사자처럼] AI 엔지니어 심화 부트캠프 (이미지 처리)

# 모두의 인테리어 Modu Interior

노혜정, 박윤수, 손명진, 차준영





[멋쟁이 사자처럼] AI 엔지니어  
심화 부트캠프 (이미지 처리)

PROJECT  
OVERVIEW

PROJECT  
RESULTS



모두의 인테리어  
Modu Interior

TECHNICAL  
REPORT

CONCLUSION



OVERVIEW

RESULTS

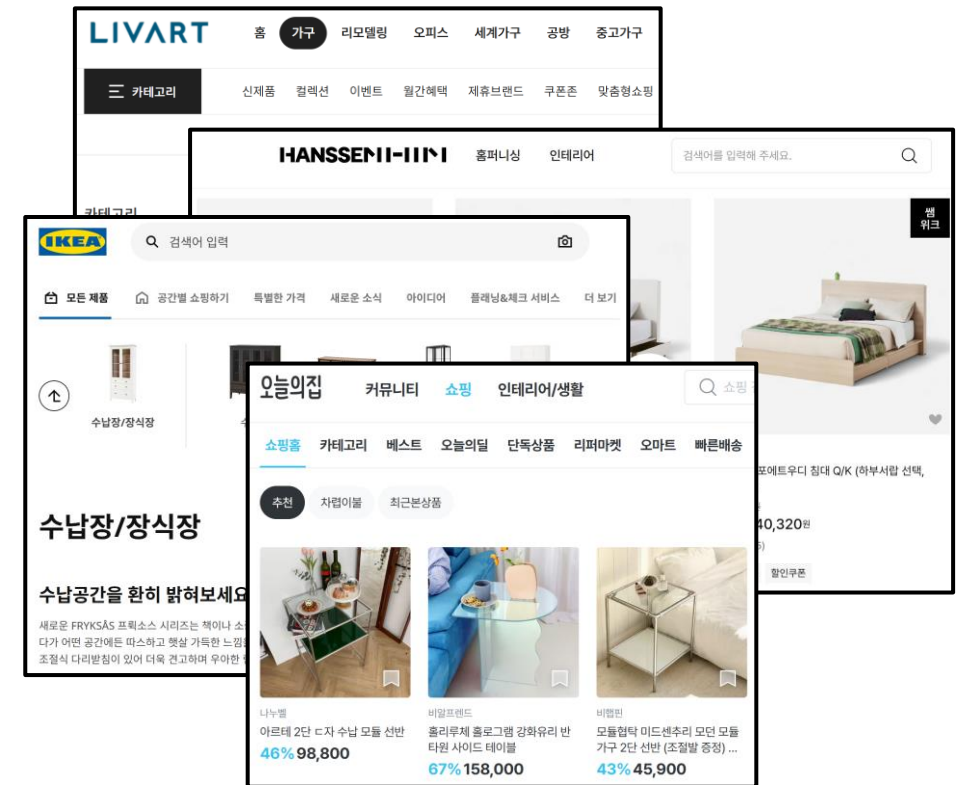
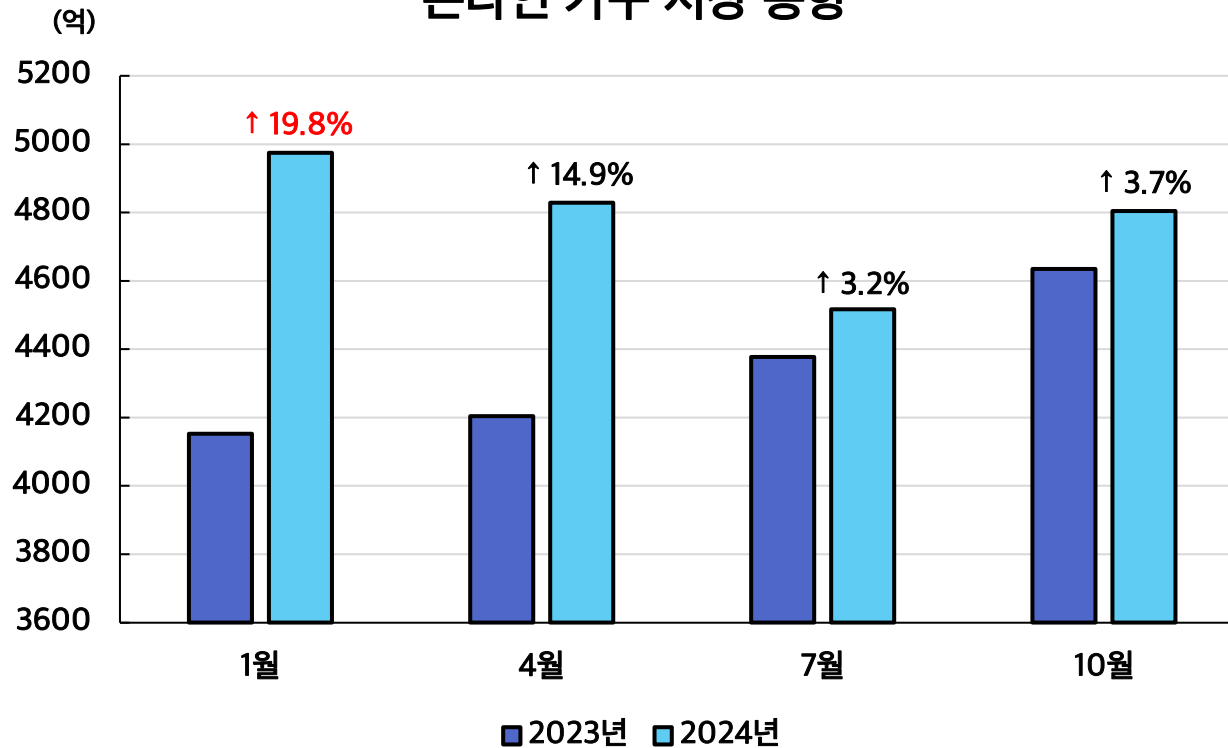
TECHNICAL REPORT

CONCLUSION

APPENDIX

## 프로젝트 배경

### 온라인 가구 시장 동향



[Fig 1.] 다양한 온라인 가구 쇼핑몰

OVERVIEW

RESULTS

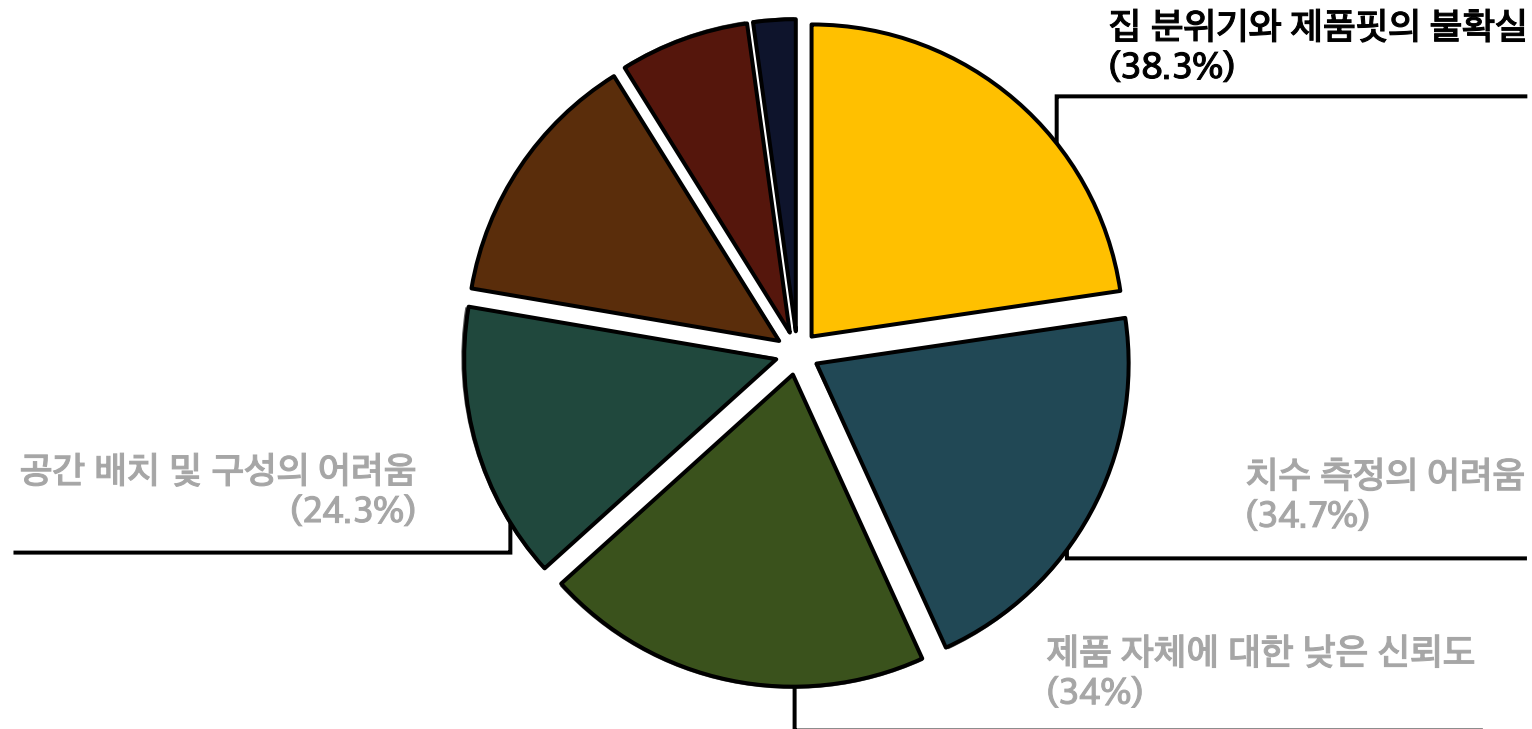
TECHNICAL REPORT

CONCLUSION

APPENDIX

## 프로젝트 배경 (Cont'd)

온라인 가구 구매 과정에서  
불편했거나 아쉬웠던 점은?





OVERVIEW

RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## 프로젝트 기대효과



### 소비자

#### 맞춤형 가구 추천

인테리어 초보자들도 자신의 공간에 잘 맞는  
가구를 찾는데 걸리는 시간을 절약할 수 있음

#### 추가 비용 손실 방지

알맞지 않은 가구 구매로 인한 교환/환불 비용  
및 재구매 시간 손실을 방지할 수 있음



### 가구 브랜드 및 플랫폼

#### 매출 증대

소비자의 가구 브랜드에 대한 신뢰도와 만족도  
를 향상 시켜 매출을 증대 시킬 수 있음

#### 중소 브랜드와 상생

가구 플랫폼을 통해 소비자의 디자인 욕구를  
만족시킬 수 있는 다양한 가구 브랜드와 상생

#### 풍부한 소비자 경험 제공

소비자가 인테리어에 참여할 수 있는 기회를  
제공하며 브랜드 마케팅으로 발전시킬 수 있음

[멋쟁이 사자처럼] AI 엔지니어  
심화 부트캠프 (이미지 처리)

PROJECT  
OVERVIEW

PROJECT  
RESULTS



모두의 인테리어  
Modu Interior

TECHNICAL  
REPORT

CONCLUSION



OVERVIEW

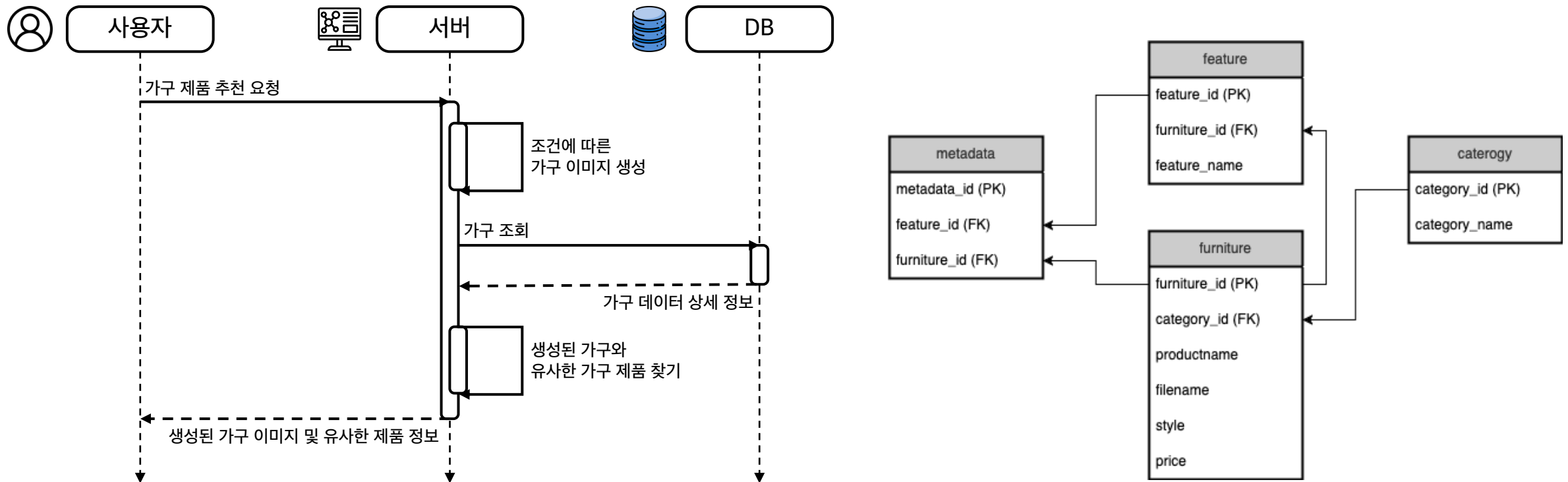
RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## 모두의 인테리어 시퀀스 차트 및 ERD



[Fig 2.] 모두의 인테리어 시퀀스 차트 및 ERD



OVERVIEW

RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## 데모 시연



[멋쟁이 사자처럼] AI 엔지니어  
심화 부트캠프 (이미지 처리)

PROJECT  
OVERVIEW

PROJECT  
RESULTS



모두의 인테리어  
Modu Interior

TECHNICAL  
REPORT

CONCLUSION



OVERVIEW

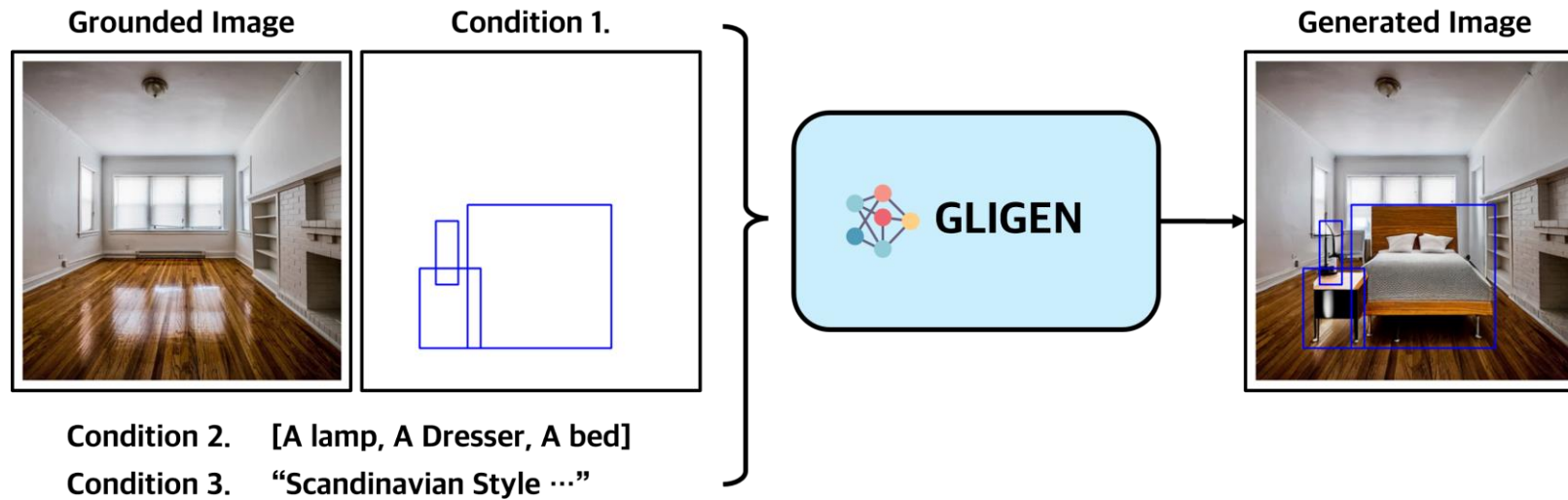
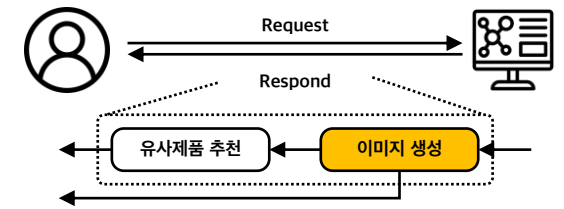
RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## Image Generation



[Fig 3.] GLIGEN based Image Generation



OVERVIEW

RESULTS

TECHNICAL REPORT

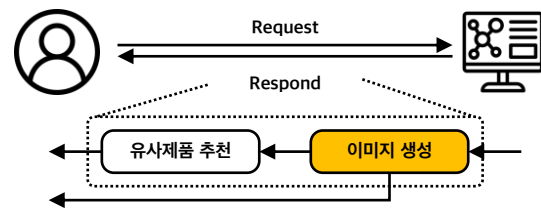
CONCLUSION

APPENDIX

## Image Generation (Cont'd)

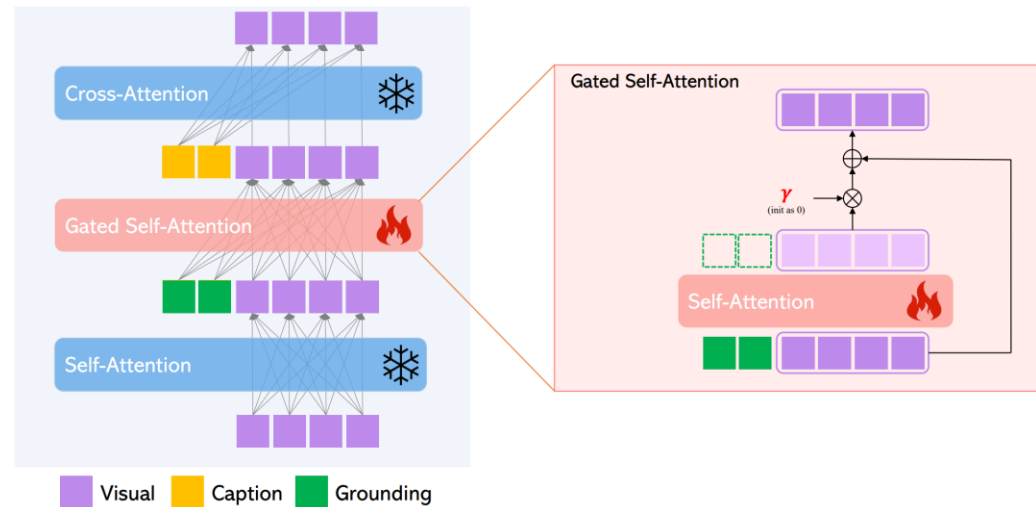
### ◆ GLIGEN (Grounded Laugange-Image Generation)

- 다양한 입력 조건을 활용해 원하는 객체를 원하는 위치에 정확히 생성할 수 있는 이미지 생성 모델
- 사전 학습 된 LDM을 기반으로 입력 조건에 따라 학습된 모듈을 추가하는 방식으로 조건에 따른 유연성이 장점



### ◆ Gated Self-attention

- Visual token의 self-attention
  - ◆ Pre-trained LDM weight, freeze
- Visual token + Grounding token의 self-attention
  - ◆ GLIGEN, trainable
  - ◆  $\gamma$ 에 따라 다양한 입력 조건에 제약을 받는 정도를 달리함
- Visual token + Caption token의 cross-attention
  - ◆ Pre-trained LDM weight, freeze



[Fig 4.] GLIGEN - Gated Self Attention



OVERVIEW

RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## Image Generation (Cont'd)

### ◆ Prompt Engineering

- 생성형 AI 모델의 학습 없이, 사용자의 입력을 최적화하여 원하는 결과를 얻는 방법

“Best quality, Cinematic lighting, Hyper-detailed textures, ...”

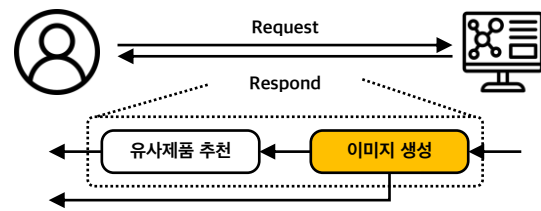
“Harmony with ambient lightening, Accurate perspective, Scale with the scene, ...”

“Scandinavian furniture design, Natural wood, Bright and airy, ...”



[Fig 5.] Multi Prompt Embedding Vectors

- 긴 프롬프트 문장을 목적에 따라 세분화하여 각각 Embedding Vector로 변환하는 방법을 사용
- 각각의 정보가 분리된 상태로 유지되어 세부적인 정보를 효과적으로 전달
  - ◆ 1. 사실적인 가구를 생성하기 위한 키워드 (Server Default)
  - ◆ 2. 방의 분위기를 고려한 가구를 생성하기 위한 키워드 (Server Default)
  - ◆ 3. **유저가 원하는 가구 스타일 키워드 (User)**



OVERVIEW

RESULTS

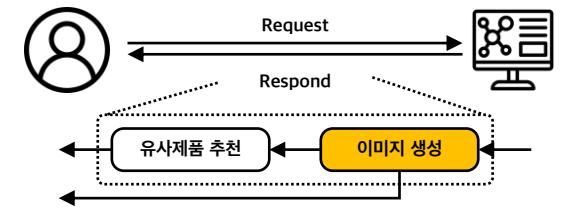
TECHNICAL REPORT

CONCLUSION

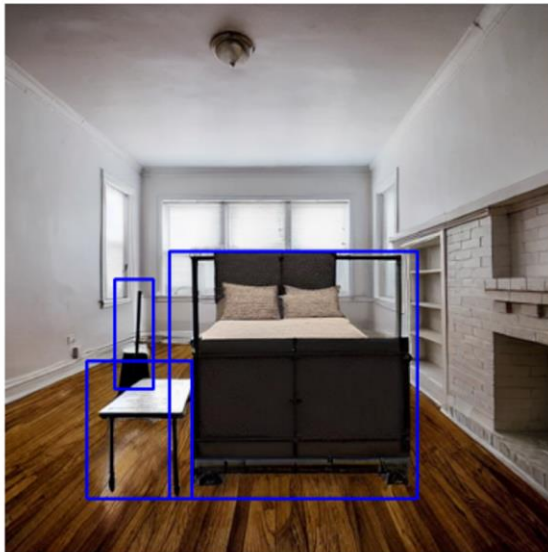
APPENDIX

## Image Generation (Cont'd)

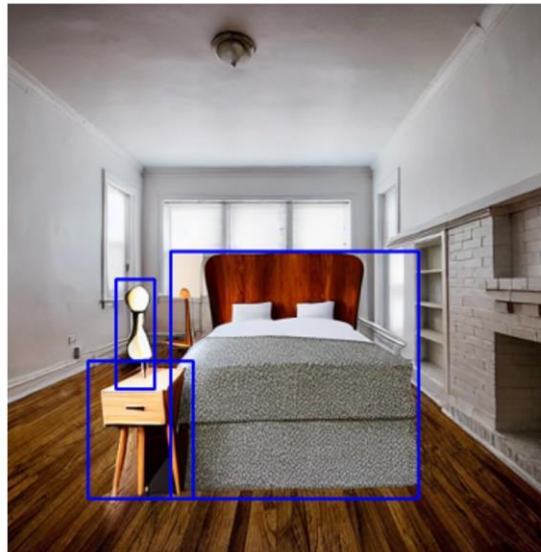
### ◆ Prompt Engineering Experimental Results



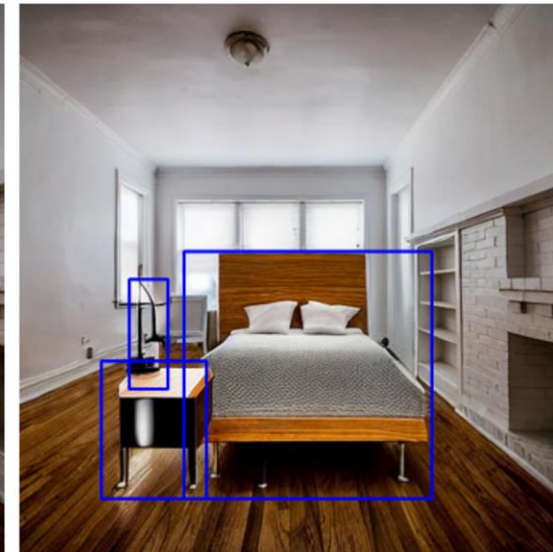
Industrial



Mid-century



Scandinavian



[Fig 6.] GLIGEN Image Generation Results

OVERVIEW

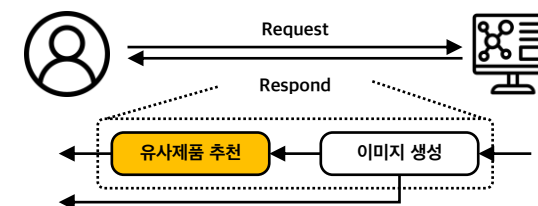
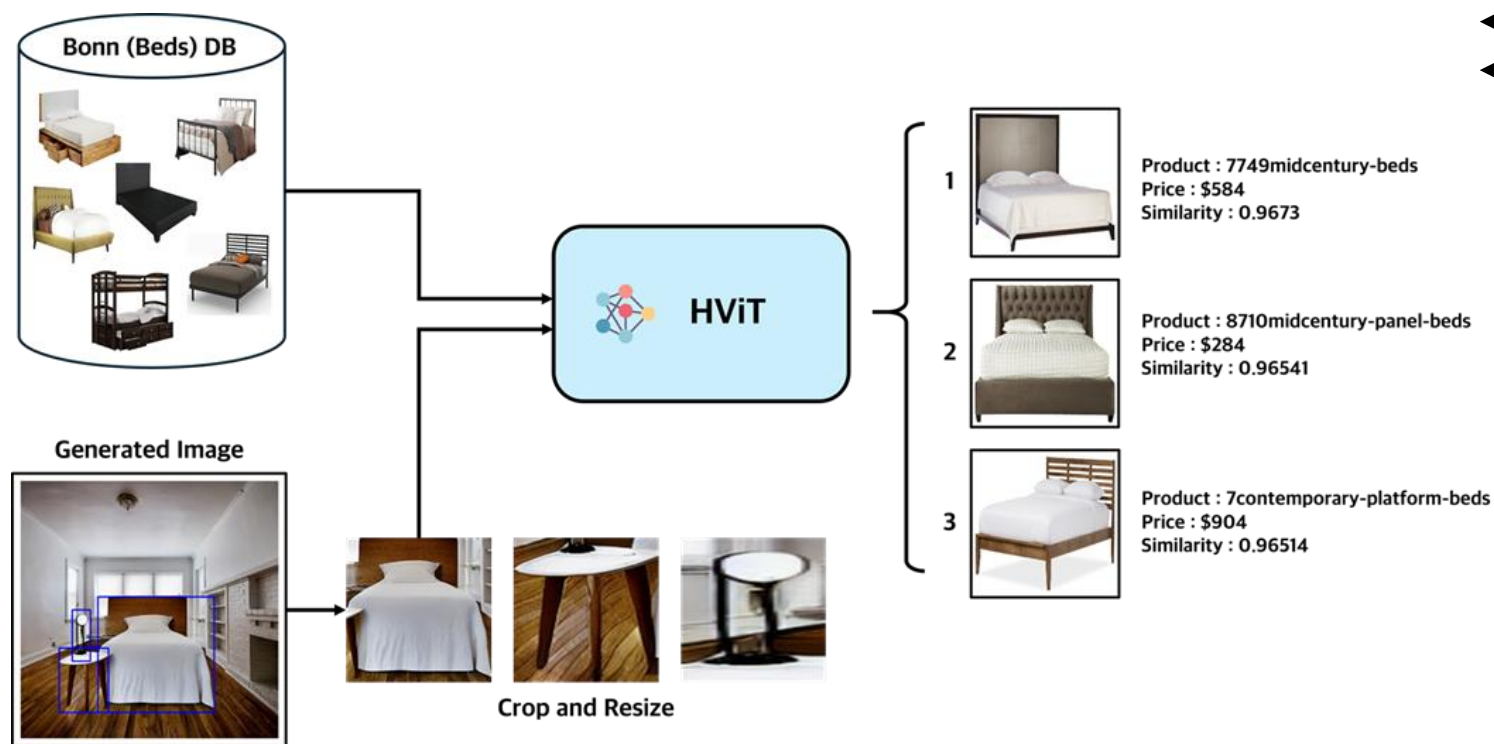
RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## Image Retrieval



[Fig 7.] HViT based Image Retrieval for Furniture Recommendation



OVERVIEW

RESULTS

TECHNICAL REPORT

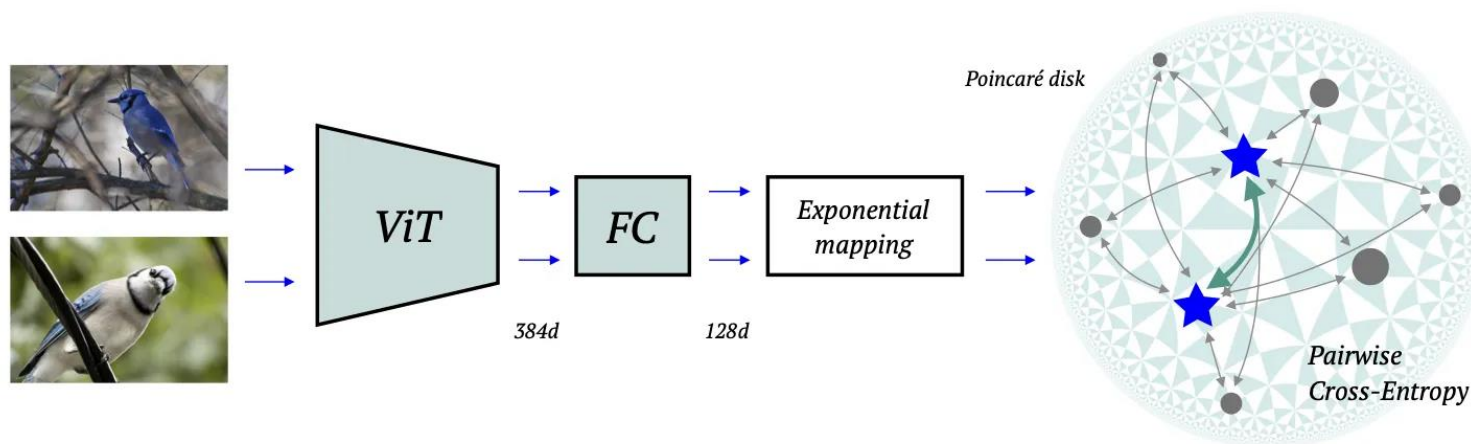
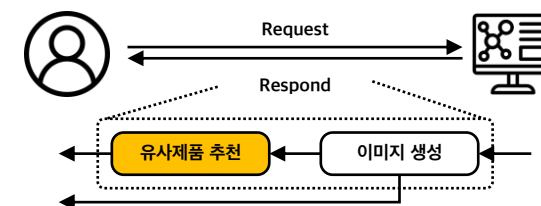
CONCLUSION

APPENDIX

## Image Retrieval (Cont'd)

### ◆ HViT (Hyperbolic Vision Transformer)

- Pair-wise Cross Entropy 손실함수로 유사한 데이터는 가까워지도록 Metric Learning
- ViT로 추출한 feature를 Poincaré Ball이라는 Non-Euclidean Space에 맵핑
  - ◆ 중심으로부터 멀어질수록 거리가 기하급수적으로 확장하는 공간의 특성을 활용해 수 많은 스타일의 데이터를 잘 구분할 수 있음



[Fig 8.] Overview of HViT Method

OVERVIEW

RESULTS

TECHNICAL REPORT

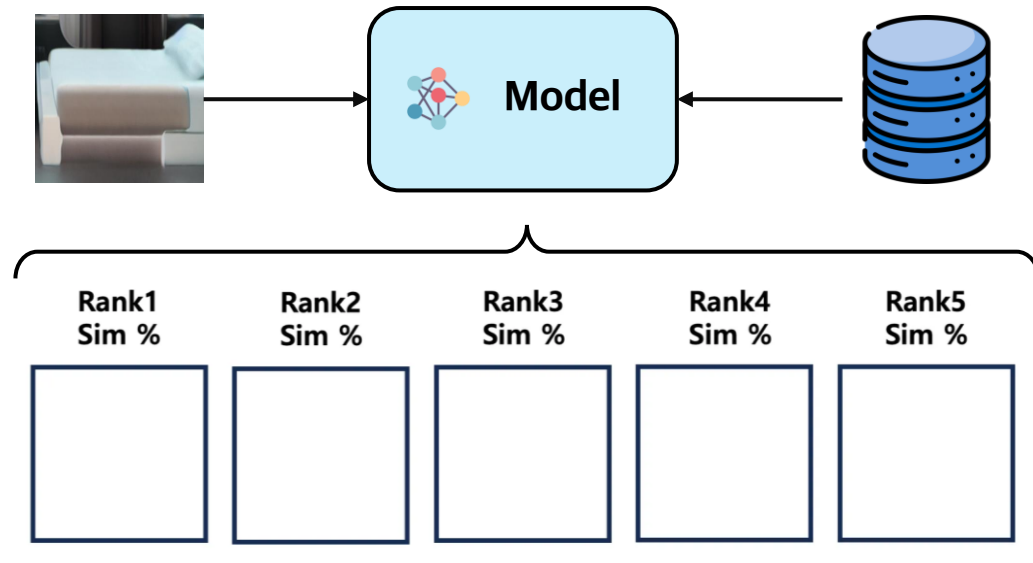
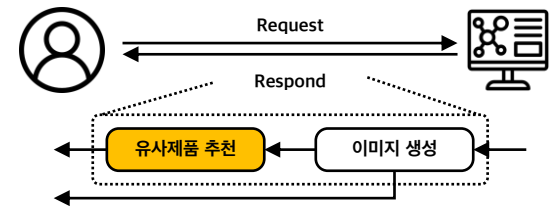
CONCLUSION

APPENDIX

## Image Retrieval (Cont'd)

### ◆ Model 별 주관적 성능 비교

- GLIGEN으로 생성된 가구 이미지로 검색 성능 테스트
- Pre-trained Model List
  - ◆ VGG16
  - ◆ ViT (Vision Transformer)
  - ◆ SimSiam (Simple Siamese Network)
  - ◆ HViT (Hyperbolic Vision Transformer)



[Fig 9.] Image Retrieval: Top-5 Similar Results

OVERVIEW

RESULTS

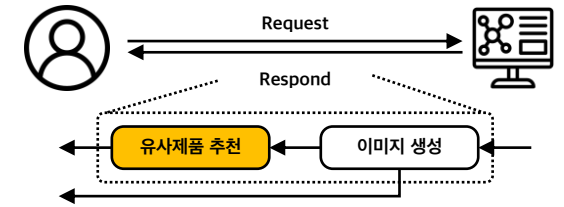
TECHNICAL REPORT

CONCLUSION

APPENDIX

## Image Retrieval (Cont'd)

### ◆ Model 별 주관적 성능 비교



(a) Generated Image

Rank 1 : 0.82639



Rank 2 : 0.82154



Rank 3 : 0.76316



(b) VGG16

Rank 1 : 0.63742



Rank 2 : 0.63719



Rank 3 : 0.62043



(c) ViT (Vision Transformer)

Rank 1  
Sim: 0.8153



Rank 2  
Sim: 0.7843



Rank 3  
Sim: 0.7789



(d) SimSiam (Simple Siamese)

Rank 1  
Sim: 0.9818



Rank 2  
Sim: 0.9817



Rank 3  
Sim: 0.9811



(e) HViT (Hyperbolic Vision Transformer)

[Fig 10.] Experimental Results: Top-3 Image Retrieval Results



[멋쟁이 사자처럼] AI 엔지니어  
심화 부트캠프 (이미지 처리)

PROJECT  
OVERVIEW

PROJECT  
RESULTS



모두의 인테리어  
Modu Interior

TECHNICAL  
REPORT

CONCLUSION



OVERVIEW

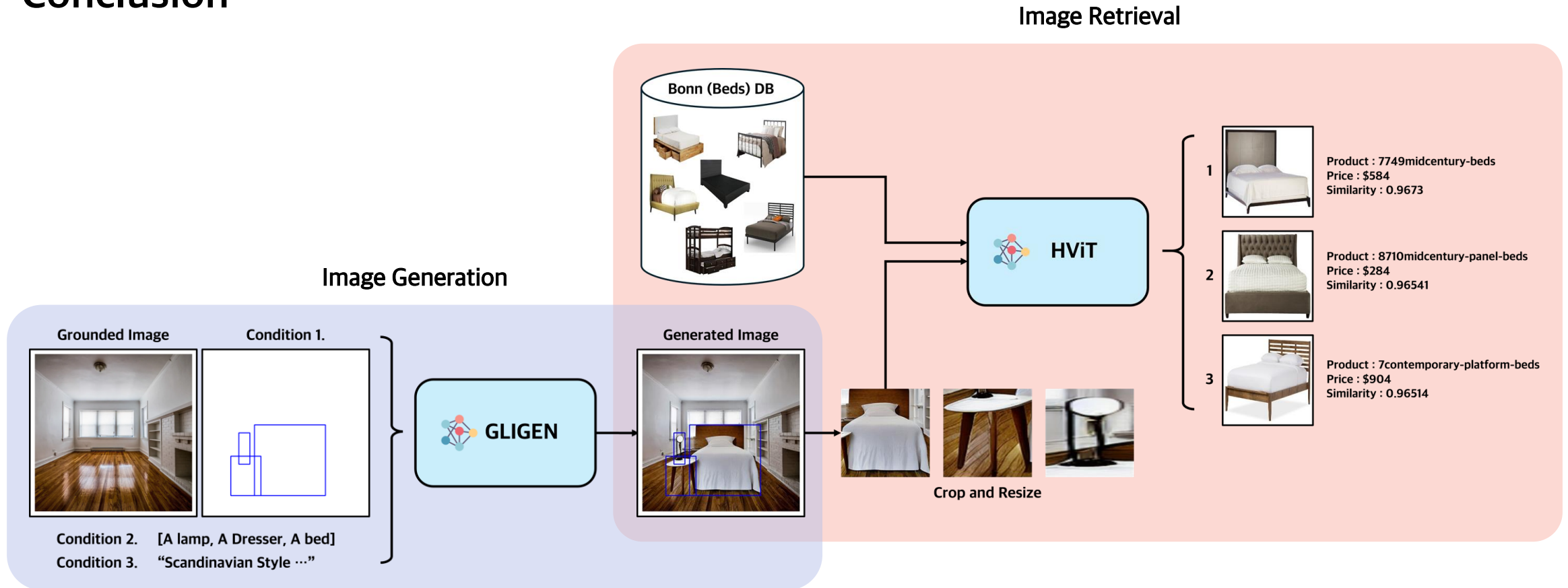
RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## Conclusion



[Fig 11.] Overview of Modu-Interior Project



OVERVIEW

RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

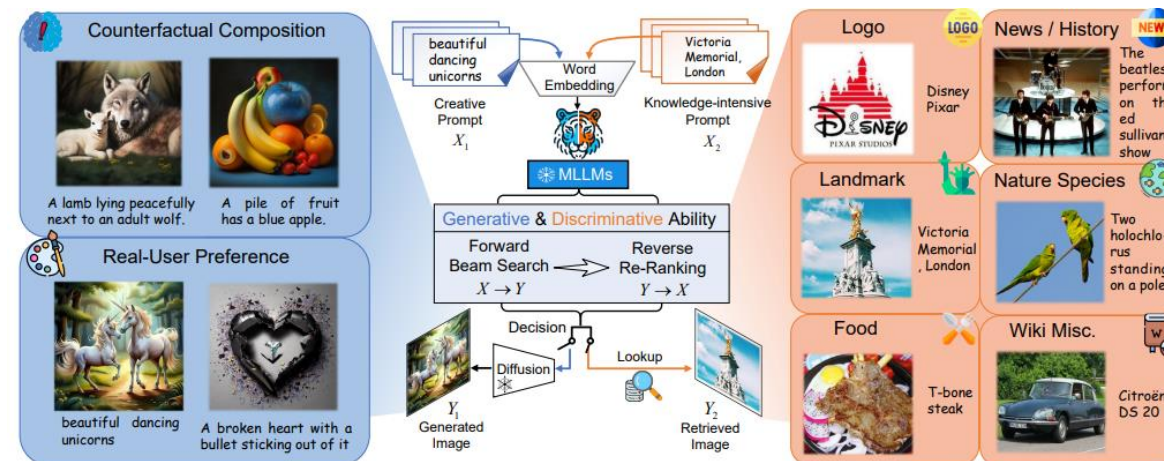
## Conclusion

### ◆ Limitations

- 원하는 가구 제품 이미지가 제대로 생성 안되는 경우
- 제대로 생성 되더라도 실제 제품이 없는 경우

### ◆ 발전 방향

- 사용자의 방에 실제로 판매되는 상품(혹은 유사한 이미지)를 생성하고 추천
- 스타일 키워드와 제품 카테고리를 입력
  - ◆ 1. 딱 맞는 제품이 있을 경우, 제품 이미지 사용
  - ◆ 2. 그렇지 않은 경우, 생성형 모델로 제품 생성



[Fig 12.] Overview of TIGer



## 팀원 소개



### 노혜정

데이터 수집, 데이터 전처리,  
웹 구현,  
스타일 맞춤 가구 생성 (이미지 생성),



### 박윤수

데이터 수집, 데이터 전처리,  
유사 제품 추천 (이미지 검색),  
객체 인식 모델 학습



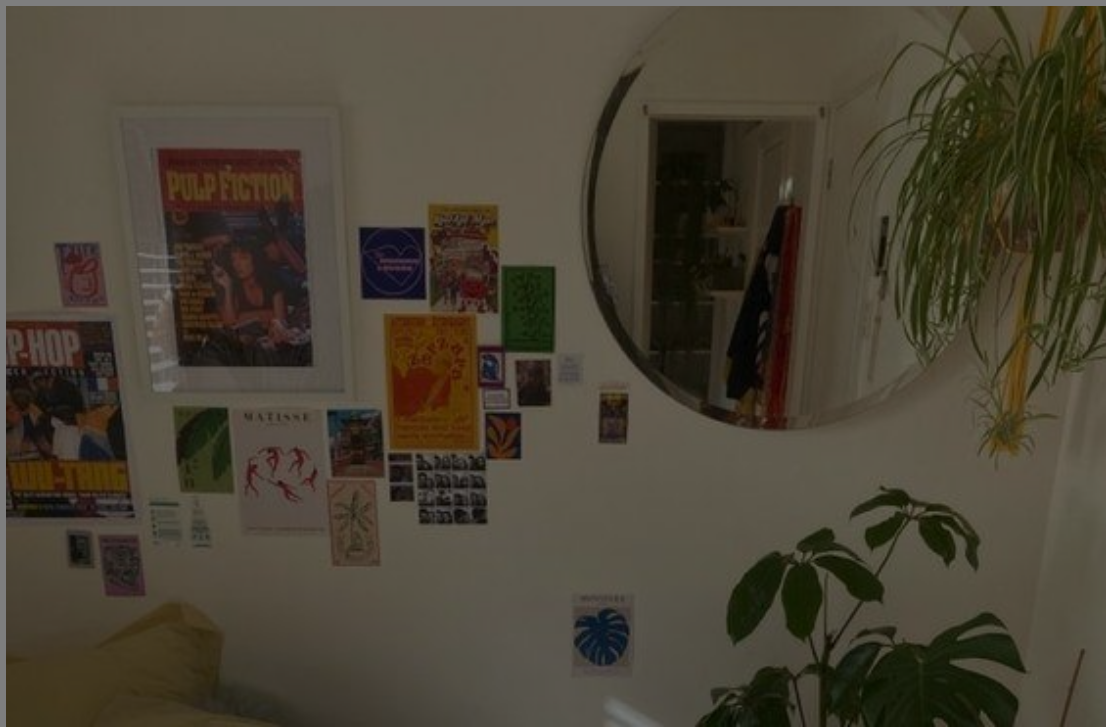
### 손명진

데이터 수집, 데이터 전처리,  
데이터 베이스 구축,  
유사 제품 추천 (이미지 검색)



### 차준영, 팀장

데이터 수집, 데이터 전처리,  
스타일 맞춤 가구 생성 (이미지 생성),  
웹 구현



[멋쟁이 사자처럼] AI 엔지니어 심화 부트캠프 (이미지 처리)

# 모두의 인테리어 Modu Interior

노혜정, 박윤수, 손명진, 차준영



OVERVIEW

RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## Appendix A. Index

### ◆ Figure

- [1. 다양한 온라인 가구 쇼핑물](#)
- [2. 모두의 인테리어 시퀀스 차트 및 ERD](#)
- [3. GLIGEN based Image Generation](#)
- [4. GLIGEN - Gated Self Attention](#)
- [5. Multi Prompt Embedding Vectors](#)
- [6. GLIGEN Image Generation Results](#)
- [7. HViT based Image Retrieval for Furniture Recommendation](#)
- [8. Overview of HViT Method](#)
- [9. Image Retrieval: Top-5 Similar Results](#)
- [10. Experimental Results: Top-3 Image Retrieval Results](#)
- [11. Overview of Modu-Interior Project](#)
- [12. Overview of TIGeR](#)



OVERVIEW

RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## Appendix B. Reference

### ◆ Paper

- GLIGEN : LI, Yuheng, et al. Gligen: Open-set grounded text-to-image generation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023. p. 22511-22521.
- Layout Diffusion : ZHENG, Guangcong, et al. Layoutdiffusion: Controllable diffusion model for layout-to-image generation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023. p. 22490-22499.
- HViT : ERMOLOV, Aleksandr, et al. Hyperbolic vision transformers: Combining improvements in metric learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022. p. 7409-7419.
- SimSiam : CHEN, Xinlei; HE, Kaiming. Exploring simple siamese representation learning. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021. p. 15750-15758.
- SEEM : ZOU, Xueyan, et al. Segment everything everywhere all at once. Advances in Neural Information Processing Systems, 2024, 36.
- ViewDIFF: HÖLLEIN, Lukas, et al. Viewdiff: 3d-consistent image generation with text-to-image models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024. p. 5043-5052.
- TIGeR : QU, Leigang, et al. Unified Text-to-Image Generation and Retrieval. arXiv preprint arXiv:2406.05814, 2024.
- Bonn Dataset : AGGARWAL, Divyansh, et al. Learning style compatibility for furniture. In: Pattern Recognition: 40th German Conference, GCPR 2018, Stuttgart, Germany, October 9-12, 2018, Proceedings 40. Springer International Publishing, 2019. p. 552-566

OVERVIEW

RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## Appendix B. Reference *(Cont'd)*

### ◆ Github

- <https://github.com/gligen/GLIGEN>
- [https://github.com/htdt/hyp\\_metric](https://github.com/htdt/hyp_metric)
- <https://github.com/facebookresearch/simsiam>

### ◆ 통계자료, 참고한 사이트

- 통계청, 2024년 1월/4월/7월/10월 온라인 쇼핑 동향
- 3D 공간데이터 플랫폼 어반베이스, 2019년 온라인 인테리어 가구/소품 구매 경험 만족도 조사
- 아키스케치 : <https://www.archisketch.com/kr/ai-interior>
- 오늘의 집 : <https://ohou.se/>

OVERVIEW

RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## Appendix C. Bonn Dataset

### ◆ Bonn Dataset

- 카테고리 : 총 6개 (beds, chairs, dressers, tables, lamps, sofas)
- 스타일 : 총 17가지 (Asian, Craftsman, Industrial, Modern, Southwestern, Tropical ... 등)
- 각 제품 이미지 별로 제조 업체, 제품 사이즈, 재질 등의 Meta 정보가 있음
- 총 90,298 장의 제품 이미지로 구성됨



[Fig C.1.] “Beds” and “chairs” from the 17 styles



OVERVIEW

RESULTS

TECHNICAL REPORT

CONCLUSION

APPENDIX

## Appendix C. Bonn Dataset (Cont'd)

### ◆ Pre-processing

- 가구 데이터셋의 품질을 낮추는 다양한 문제가 발생함



- 데이터셋의 일관성을 높이고 검색 기능의 최적화를 위해 다음 방법에 따라 DB 구축
  - ◆ 1. 대중적으로 많이 선호하는 5가지 스타일 선택 (Contemporary, Industrial, Modern, Mid-century, Scandinavian)
  - ◆ 2. 각 스타일 별로 최대 50장의 좋은 퀄리티 제품 이미지 선정

OVERVIEW

RESULTS

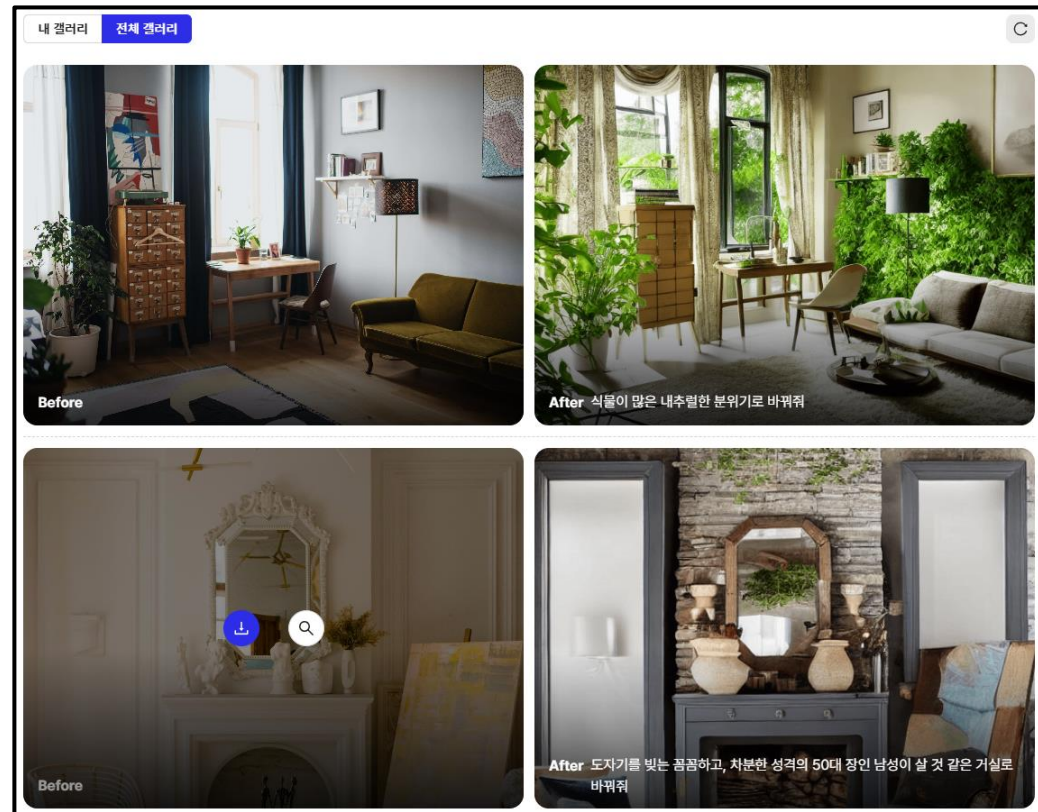
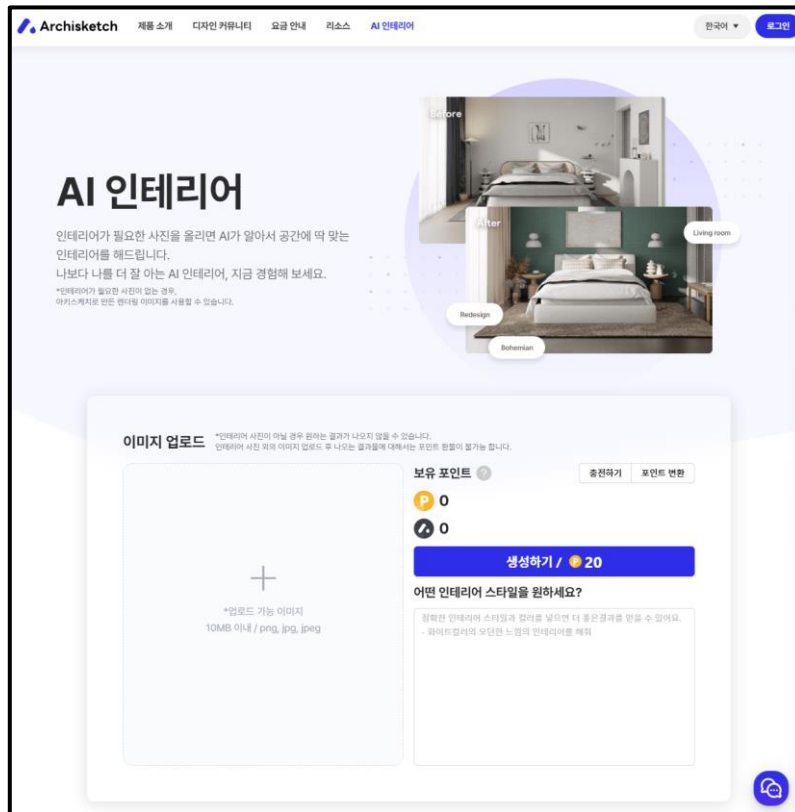
TECHNICAL REPORT

CONCLUSION

APPENDIX

## Appendix D. 유사 서비스와 비교

### ◆ 아키스케치 : Archisketch



[Fig D.1.] 아키스케치 서비스

OVERVIEW

RESULTS

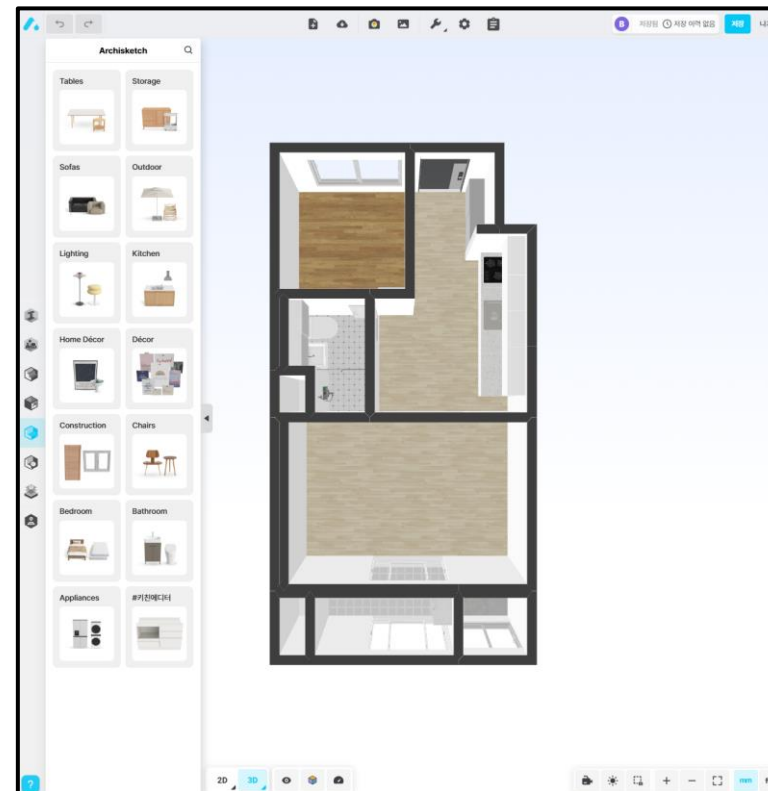
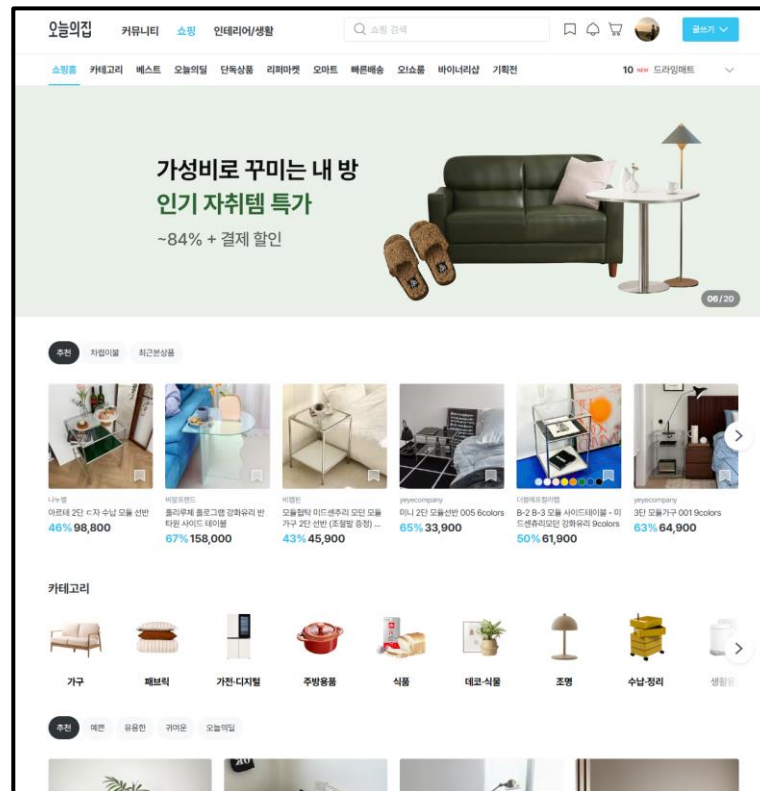
TECHNICAL REPORT

CONCLUSION

APPENDIX

## Appendix D. 유사 서비스와 비교 (Cont'd)

### ◆ 오늘의 집



[Fig D.2.] 오늘의 집 서비스