



Copybara

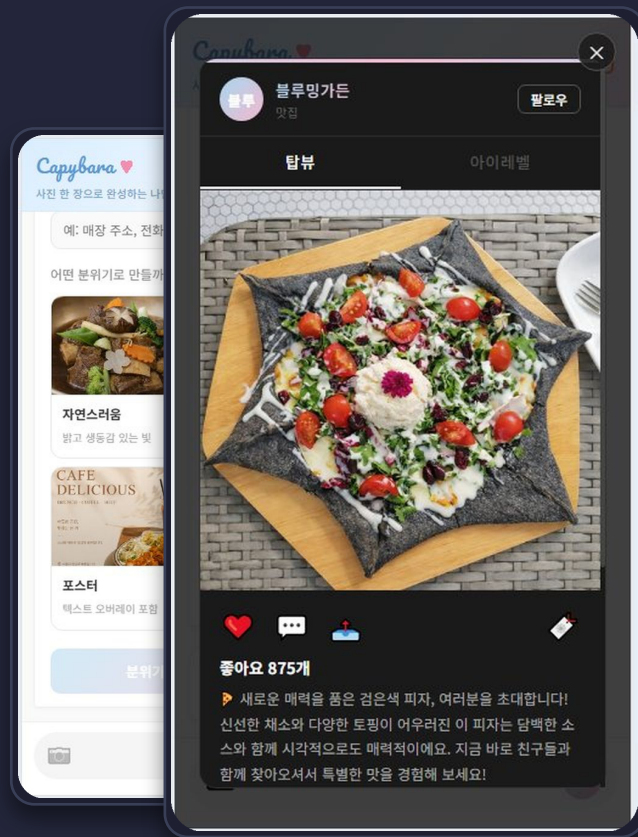
카피바라

음식 사진 한 장이면, 2분컷 인스타 피드 완성

AI를 몰라도 쉽게 쓰는 외식업 광고 콘텐츠 생성기

코드잇 AI 부트캠프 7기 · 3팀 카피바라

app.copybara.p-e.kr



발표 순서 & 팀 역할

발표 순서 Agenda

- | | | |
|---|-----------------|----|
| 1 | 도입 · 솔루션 | 윤민 |
| 2 | 데모 · 차별성 · 동작원리 | 봄 |
| 3 | 데이터 · 모델 선정 | 지훈 |
| 4 | OpenAI 심화 · 평가 | 민선 |
| 5 | HF SDXL · 모델 병행 | 정우 |
| 6 | 인프라 & 배포 & 운영 | 가빈 |
| 7 | 성과 · 한계 | 윤민 |

팀 역할 분배 Team

 봄 PM · Backend	 윤민 PM · Frontend
 지훈 Model · 검수	 민선 Model · OpenAI
 정우 Model · HF	 가빈 Infra · 배포 · 운영

외식업 소상공인의 3 No

“감각 없음 · 시간 없음 · 기술력 없음” — 광고를 만들고 싶어도 시작조차 막힌다.



감각 없음

디자인 도구·광고 문구 작성이
막막하다



시간 없음

꾸준한 SNS 운영에 들일
시간이 없다



기술력 없음

사진 보정·피드 제작 반복이
버겁다



주 타겟 온라인·SNS에 익숙치 않은 30~50대 외식업 주 운영층

왜 외식업에 집중했나

51%

소상공인 중 외식업 비중
외식·카페·베이커리
종사자

50%

홍보 채널 인스타 활용률
비주얼 마케팅 의존도
특히 높음



타겟을 좁혀 페르소나를 명확히

외식업

SNS 유입이 곧 매출 · 감성/MZ 분위기

뷰티·네일

포트폴리오·시술 전후 중심

펫·동물병원

귀여운 콘텐츠 바이럴

동네 의류·잡화

상세페이지 취약, 전환형

→ SNS 의존도가 가장 큰 외식업을 선택

사진 한 장 → 인스타 피드 완성

1. 가게 이름

2. 사진 업로드

3. AI 분석 확인

4. 옵션 선택

5. 피드 완성

③ 분석 결과 '맞아요 / 아니에요' 확인 후 진행 ⑤ 산출물 탭뷰·아이레벨 2종 + 본문·해시태그

핵심 기능 3가지



가이드형 UX

자유 대화 없이 버튼·카드만으로 진행. 다음 할 일을 항상 안내해 이탈을 막는다.



이미지 생성 옵션
4가지

자연스러움·감성적·포스터·일러스트
4종 중 원하는 분위기를 골라
같은 사진을 다르게 연출한다.



분석 기반 맞춤 생성

음식을 분석해 카테고리·특징에
맞는 이미지·문구를
만든다.

백 마디보다 한 번, 약 1분 데모

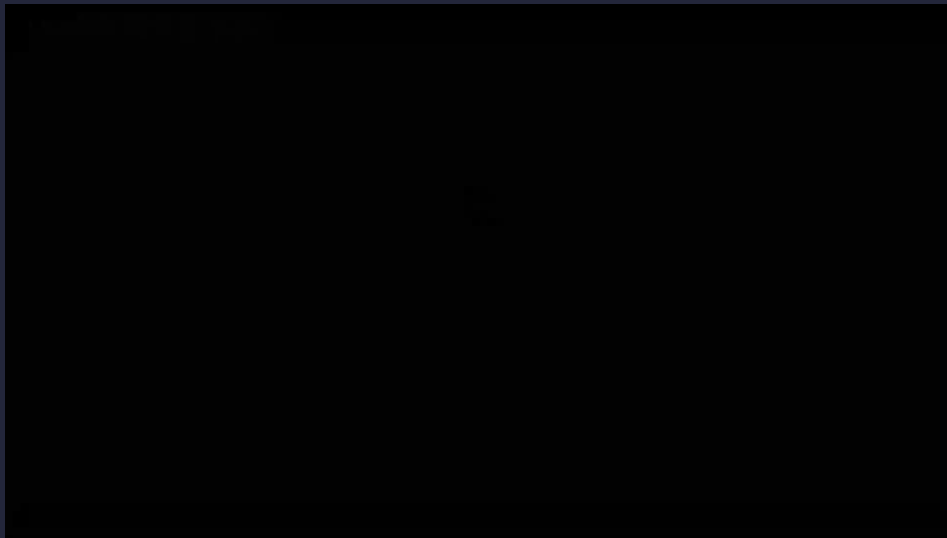


사진 한 장으로 완성되는 인스타 피드 콘텐츠

같은 사진, 같은 요청 → 다른 결과

입력



인스타 올릴 광고 피드용으로 만들어 줘

원본 음식 사진 + '인스타 광고 피드용으로 만들어 줘'

카피바라



인스타 피드 완성형 — 탭뷰/아이레벨 2종 + 본문 · 해시태그

ChatGPT



포스터 1장 — 정보 플레이스홀더(@your_cafe)

Gemini



포스터 1장 — 가게명 · CTA 임의 생성(Café Celestial)

Copybara의 차별점

① 포스터 1장 → 피드 완성형

② 정보 정확성 = 환각 방지

③ 결과 통제·일관성

	카피바라	ChatGPT	Gemini
산출물 형태	인스타 피드 완성형	포스터 1장	포스터 1장
이미지 구성	탐뷰·아이레벨 2종	단일 이미지	단일 이미지
본문·해시태그	○ 분리 제공	× 이미지에 박힘	× 이미지에 박힘
감성·구도 선택	○ 4종 × 2구도	× 불가	× 불가
가게 정보	○ 사용자 입력 기반	△ 플레이스홀더 임의	× 가게명 지어냄
바로 업로드	◎ 그대로 가능	△ 수정 필요	△ 허위정보 수정

Copybara는 음식 사진 한 장에서 → 소상공인이 바로 쓸 수 있는 광고 콘텐츠까지 연결한다.

외식 특화 × 무기술 × 자동 생성 = 카피바라뿐

↑ 외식 피드 이미지 자동 생성·특화



즉시성·셀프 자동화 (낮음 → 높음)

교집합은 카피바라뿐

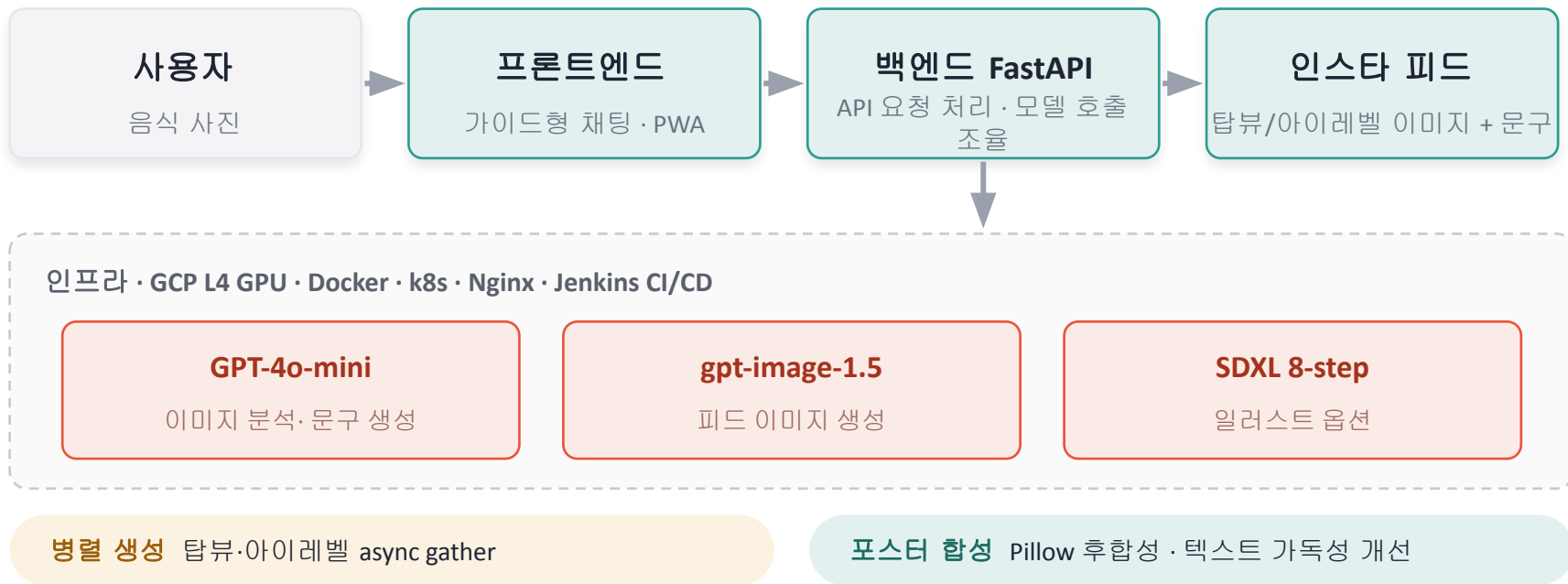
- ✓ 외식 특화
- ✓ 무기술 (선택만)
- ✓ 피드 이미지 자동 생성

Canva·미리캔버스 직접 편집 필요, 외식 특화 아님

크몽 외주 사람이 제작, 고가·느림

킵그로우 계정 운영 대행, 피드 자동생성 X

시스템 아키텍처



분석 체계 & 모델 선정

서비스 파이프라인

음식 사진 1장

① 이미지 분석

② 광고 문구 생성

③ 이미지 생성



카테고리 & 분석 속성

카테고리 (5) 한식 · 중식 · 일식 · 양식 · 기타

분석 속성 (6 필드) category · product_name · main_features · color · mood · target



Ground Truth — 하이브리드

사실 필드 (category · product_name)

→ 이미지 메타데이터 CSV에서 수동 정의

서술 필드 (main_features · color · mood · target)

→ GPT-4.1-mini가 이미지를 직접 보고 생성

모델 선정

단계	선정 모델	비교 대상	선정 이유
이미지 분석	gpt-4o-mini	gpt-4.1-mini · gpt-5.4-nano	정확도·속도·비용 균형이 우수
문구 생성	gpt-4o-mini	gpt-4.1-mini · gpt-5.4-nano	키워드 반영·톤 유지가 안정적
이미지 생성	gpt-image-1.5	gpt-image-1-mini	프롬프트 반영력·이미지 완성도 높음

평가 데이터셋 & 지표



평가 데이터셋

카테고리별 20장 × 5 = **100장** 수집 → API 비용 고려해 **30장 subset**으로 메인 평가

단계	평가 데이터	평가 건수	비고
Analyze	subset 30장	30건	
Copy	subset 30장 × 3 tones	90건	친근·고급·유머
Generate	subset 30장	30건	

분석·문구 단계는 전체 **100장**으로 일반화 성능도 검토 (G-Eval·이미지생성은 30장)

단계	지표	실제 사용 데이터
이미지 분석	attribute_f1 · clip_score	100장
문구 생성	keyword_coverage · tone_similarity	100장 × 3 = 300건
문구 생성	G-Eval	30장 × 3 = 90건
이미지 생성	identity · aesthetic · clip_iqa	30장



평가 지표

목적 주관 판단 대신 각 단계 품질을 객관 측정 → 개선방향을 수치로 검증

단계	평가축	지표
이미지 분석	속성 추출 정확도	속성 단위 F1 score
	텍스트-이미지 정합성	CLIPScore
광고문구 생성	톤 일치성	임베딩 코사인 유사도
	자연스러움·매력도	G-Eval (gpt-4.1-mini)
	사용자 입력 반영도	키워드 커버리지
이미지 생성	음식 정체성 보존	DINOv2 코사인 유사도
	광고 품질 향상	LAION Aesthetic (Δ)
	광고 품질 향상	CLIP-IQA (Δ)

이미지 생성 품질 개선

문제 과도한 AI 느낌 · 음식 정체성 과도 변형

자영업자가 찍은 음식의 정체성을 보존한 채 광고용으로 미적 개선이 목표 → 원본 보존이

원인 gpt-image-1-mini · generate 방식 · 분석 텍스트만 사용

→ 음식 정체성 보존의 어려움

해결 4가지

1 모델 변경

gpt-image-1-mini → gpt-image-1.5

2 프롬프트 고도화

- 1) 보존·변경 대상 분리
- 2) food포토그래퍼 역할부여

3 input_fidelity

원본 이미지 반영 강도 조절

4 generate → edit

- 1) 원본이미지 직접 입력
- 2) 텍스트는 보조정보

× 베이스라인 생성 — 변형·AI 느낌



✓ 개선 결과



결론 원본 음식 정체성 보존 + 광고 이미지로서의 품질 개선을 동시에 달성

이미지 생성 옵션별 결과

사용자 니즈에 맞게 4가지 스타일 제공 — 자연스러움 · 감성적 · 포스터(OpenAI) · 일러스트(HF)

● 자연스러움

음식·배경은 원본 그대로,
조명·색보정·초점 등 촬영 품질만 개선



● 감성적

음식 정체성은 유지하되,
배경·분위기 재구성 · 소품 추가 허용



● 포스터

텍스트 오버레이용 여백 설계 ·
음식 우하단 고정 (배경 생성이 목표)
pillow 후합성



Pillow gpt가 이미지를 생성한 뒤, 그 이미지 위에 텍스트를 코드로 직접 그려 엮는 것. 텍스트는 파라미터를 재활용.

베이스라인 대비 항목별 개선 결과

subset 30장 기준 · 이미지 분석 / 문구 생성 / 이미지 생성 3개 항목 정량 측정(메인평가)



이미지 분석

속성 F1 0.455 → **0.554**

CLIPScore 0.524 → **0.549**

개선사항

- 임베딩 유사도 기반 soft matching 도입
- 색상·분위기 어휘 16종 표준화



문구 생성

G-Eval 총합 4.01 → **4.30**

톤 일치 3.51 → **4.54**

키워드 커버리지 0.922 → **0.915**

개선사항

- 톤별 어휘 가이드 프롬프트 주입
- 메인·서브·해시태그 구조화 포맷 고정



이미지 생성

DINOv2 유사도 0.740 → **0.872**

정체성 보존율 0.70 → **0.95**

개선사항

- 분위기별 베이스 프롬프트 분리
- 유지·재구성·금지 제약조건 세분화

※ 자연스러움 기준 · 감성/포스터도 보존율 0.9+



핵심 가장 약했던 톤 일치 +29.3% · 정체성 보존 +35.7% 로 크게 향상



트레이드오프 키워드 커버리지 -0.8%(구조화 출력포맷 고정) · 포스터 Aesthetic 하락(텍스트 합성용 여백 확보)

추가 정량평가 — subset 대표성 확인



목적 개선모델의 안정성 및 일반화된 성능을 보완적으로 검토(보조평가)

지표	베이스라인	개선후 (subset 30장)	개선후 (전체 100장)	30 vs 100
속성 F1	0.455	0.554 (+21.8%)	0.527 (+15.8%)	100장 ↓
CLIPScore	0.524	0.549 (+4.8%)	0.570 (+8.7%)	100장 ↑
키워드 커버리지	0.922	0.915 (-0.8%)	0.885 (-4.1%)	100장 ↓↓
톤 유사도	0.411	0.428 (+4.3%)	0.425 (+3.4%)	거의 일치



확장평가

개선 모델이 더 큰 평가셋에서도 전반적인 개선 경향을 유지하는지 확인한 보조 평가로 해석

↔ 트레이드오프

키워드 커버리지 하락폭(-4.1%)이 표본(-0.8%)보다 뚜렷
 → 톤별 어휘 가이드 · 구조화 포맷 → 키워드 수용 유연성 감소
 → 톤 일치 개선 과정에서 발생한 구조적 트레이드오프로 해석

파이프라인 구축 및 고도화 과정 *Pipeline Evolution*

시도 & 한계 · 실패 반복



txt2img

카테고리별 배경 프롬프트로 이미지 생성

결과 원본 음식 사진이 전혀 반영되지 않음

↓ 원본 보조 필요 → *img2img* 전환



img2img

원본에 노이즈 추가 후 스타일 변환

결과 스타일 변환은 성공, 배경만 따로 교체 불가

↓ 배경 분리 필요 → *rembg* 시도



rembg + inpainting

음식 자동 추출 후 배경 교체

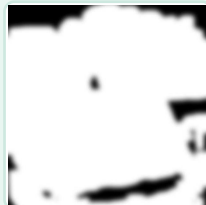
결과 철판·반찬 많은 이미지에서 음식 추출 실패

✓ 최종 채택

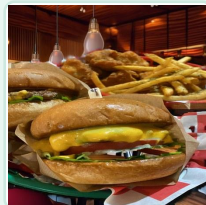
Grounded SAM

텍스트 기반 의미론적 감지로 기존 한계를 극복

Step 1



Step 2



Step 3



- **Step 1 · Grounding DINO + SAM** 'food·pot·bowl' 등 프롬프트 → 픽셀 단위 마스크 생성
- **Step 2 · SDXL Inpainting** 마스크 기반으로 배경 영역만 새로운 배경으로 교체
- **Step 3 · SDXL img2img** 전체 이미지를 팜아트·애니(서비스 내 일러스트 옵션으로 반영) 스타일로 변환

최종 배치 테스트 성능 평가 및 향후 기술 로드맵

Performance & Roadmap

오픈소스 모델 실증 완료 — 서비스 고도화를 위한 차기 과제

파이프라인 성능 평가 Batch Test

정량 성과 · 오픈소스 모델 가능성 실증

100%

음식 감지 성공률
8 / 8 전체 성공

31.7초

평균 처리 속도
JIT·이상치 제외

20.94GB

최대 GPU VRAM
Step 3 완료 시점

정성 평가 & 한계

Ad-Ready 62.5%

✓ 단순 접시·그릇 구조 마라탕·초밥·라멘·파스타·버거 등 높은 완성도

! 복잡한 용기 구조 순두부찌개(반찬)·닭갈비(철판)
·담성(찜통)에서 마스크 반전·구멍 문제



담성 사례

차기 버전 개선 과제 Roadmap v2.0

1

마스크 품질 개선 · 자동화

fill_bbox=True로 철판·찜통 내부 마스크 구멍 해결 · max_coverage 임계값으로 과마스킹 방지

2

광고 스타일 다양화 (LoRA)

학습된 LoRA(food_moody, scale 0.6)로 SDXL 의존 탈피 → 'Dark & Moody' 등 B2B 비주얼 확장

3

실서비스 속도 최적화

경량 SAM ViT-B(375MB) + DINO 배치 추론 → 응답 16초 이내 단축 목표

4

IP-Adapter 도입 — 원본 참조 강화

한계 · 생성(Step 2·3)이 텍스트에만 의존 → 색감·질감·구조 반영 미흡

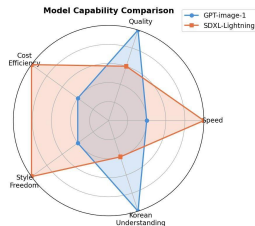
방안 · CLIP 이미지 인코더 IP-Adapter를 Step 3에 컨디셔닝 주입

효과 · GPT·Gemini 수준 원본 충실도 + 스타일·보존 동시 향상

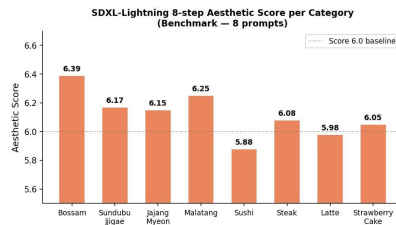
이미지 생성 · 모델 비교

GPT-image-1.5 & SDXL-Lightning

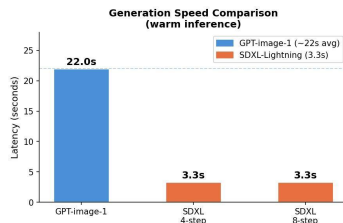
● 모델 역량 비교



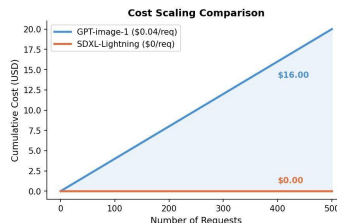
● SDXL aesthetic



● 생성 속도 비교



● 요청량에 따른 누적 비용



서비스 내 역할 분담

스탠다드 → **GPT-image-1.5**

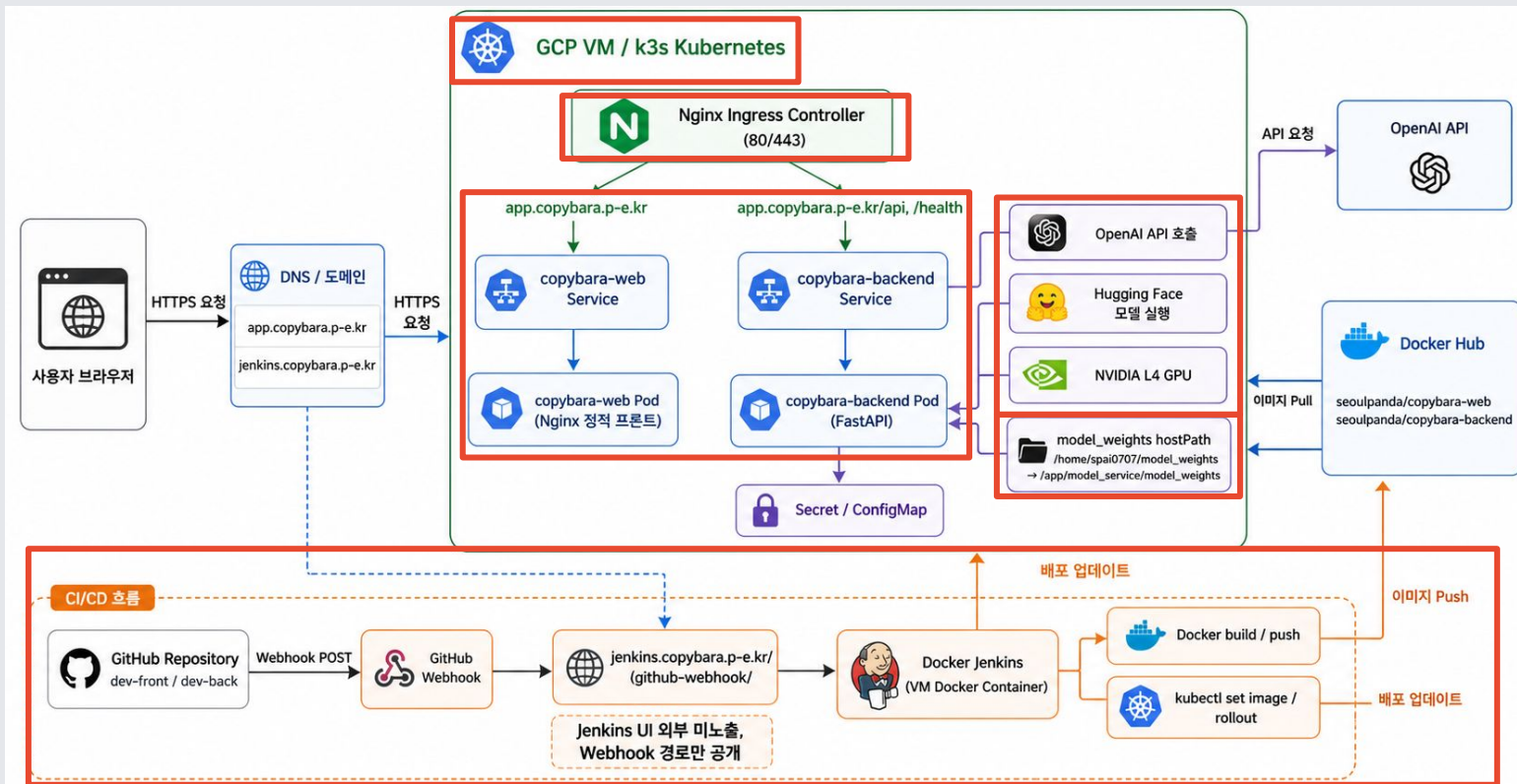
자연스러움 · 감성적 · 포스터 → 고품질 · 한국어 직접 처리 · 안정적

일러스트 → **SDXL-Lightning**

팍아트/애니 · 3.3초 · 비용 없음 · 스타일 자유 → 차별화 옵션 · 대량 저비용 운영

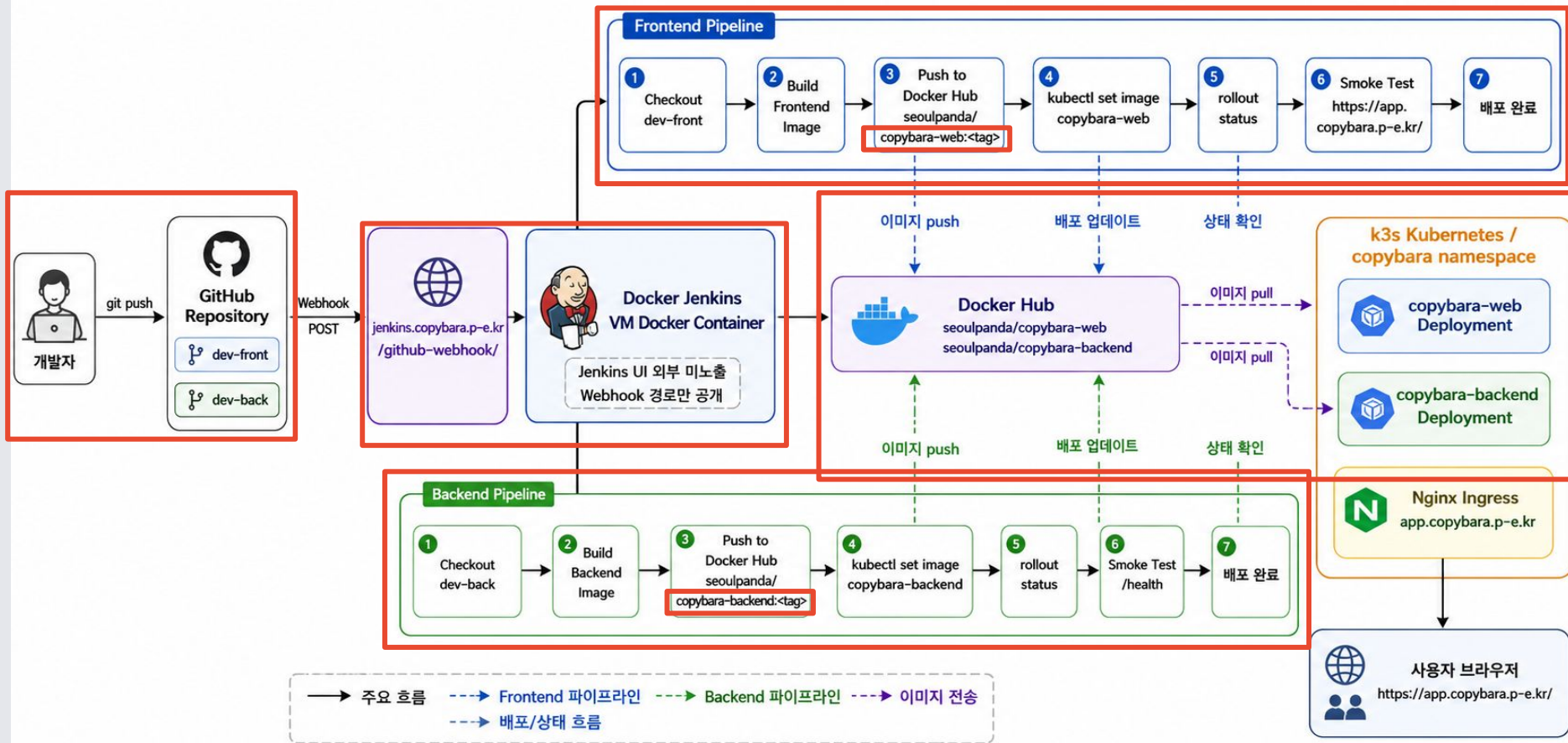
두 모델은 경쟁 관계가 아닙니다. GPT는 품질과 안정성이 필요한 기본 서비스에, SDXL은 속도 · 비용 · 스타일 확장이 필요한 곳에 — 각 강점을 살린 구성으로 두 모델을 **함께 활용**합니다.

Copybara 전체 인프라 아키텍처



Copybara CI/CD 흐름

발표 · 가빈님



자동 배포 파이프라인



GCP VM · L4 GPU

모델 서빙용 자체 GPU 환경



Jenkins ↔ GitHub 웹훅

푸시 시 자동 빌드·배포



Kuburnetes로 Pod관리

배포 장애 시에도 죽지 않는 서버

핵심 성과

M1 최소 기능 · 업로드→분석→생성 ✓

M2 고도화 · 카테고리·톤·해시태그·언어 ✓

M3 옵션 4종 · 2종 결과 · UI/UX · 배포 ✓

영역 · 지표	Before	After	비고
이미지 분석 · 속성 F1	0.455	0.554	Precision 0.53→0.64
광고 문구 · 톤 일치	3.51	4.54	+29.3%
광고 문구 · G-Eval 총합	4.01	4.30	커버리지 91.5%
이미지 생성 · 정체성 보존율	70%	88~95%	edit 전환
이미지 생성 · Aesthetic Δ	—	+0.20~+0.50	포스터 외 개선

핵심 가장 약했던 톤 일치 +29.3%, 정체성 보존 70→최대 95%로 최대 개선

달성 요약 — 핵심 성과 5선

- 정체성 보존 파이프라인 확립**
generate→edit + input_fidelity → 70%→88~95%
- 톤 반영형 문구 생성**
어휘 가이드 + 출력 구조 고정 → 톤 일치 +29.3%
- 5단계 가이드형 UX 완성**
자유대화형→가이드형, 사용자 이탈 최소화
- GPT·SDXL 이중 모델 병행**
품질·한국어 + 비용 \$0·스타일, 음식 감지 100%
- 플스택 서비스 배포 (웹·앱 모두 지원)**
HTTPS·PWA·CI/CD·GPU 운영, 마일스톤 100%

향후 과제 — Known Issues & Roadmap

인지·대응 중인 과제부터 백로그까지 — 4개 개선 트랙



1. 품질 · 성능 개선

대응 과제

- AI 분석·생성 정확도 고도화 — 오분류 사례 → 프롬프트 출력 형식 예시 구체화(개선안 작성), 모델 고도화로 포스터 옵션 이미지 성능↑
- 이미지 생성 시간 단축 — `asyncio` 병렬 적용, 추가 단축은 스트리밍 등 구조 변경(후속 과제)
- 일러스트 옵션 — '아이레벨뷰' 생성 이미지 추가



2. 기능 확장

다음 버전

- 포스터에 사용자 추가 정보(주소·전화·이벤트·신규 할인) 반영
- 가게 이름 검색(카카오로컬 API) 연동
- 공유 기능 R2 URL 기반 전환 — 인프라팀 R2 연동 완료 후
- PWA 앱 설치(다운로드) 유도 UI 웹페이지 내 추가



3. 개인화 · 사업화 · 인프라

다음 버전

- 로그인+계정 기반 커스터마이징 — 브랜드 등록, 자주 쓰는 스타일·포맷 저장
- 유료화 시스템 — 무료 체험 / 유료 생성 횟수 구분
- Prometheus/Grafana 기반 리소스 및 장애 모니터링
- CPU/Memory 기반 자동 스케일링 구성
- k6 등을 활용한 대량 트래픽 테스트



4. 구조 개선

리팩토링 · 백로그

- 단계 기반 렌더링 전환 — `renderStepN()+currentStep`으로 단순화(완료=정적 말풍선, 현재만 인터랙티브) → 유지보수성↑
- 백엔드 API 서버와 모델 추론 서버 분리
- 프로필 아바타 AI 썸네일 생성 (백로그)



THANK YOU

AI를 몰라도, 누구나 쉽게.

음식 사진 한 장으로 완성하는
외식업 광고 콘텐츠 생성 서비스 — 카피바라

app.copybara.p-e.kr

