

사람을 대신할 답이 아니라 사람에게 확인할 질문을 만든다

합성 페르소나 연구는 미래 시나리오와 정책 가설의 점검 도구를 만든다.

지금 쓸 곳

모호한 문항, 수용 조건, 반대논리와 후속 질문 후보를
체계적으로 점검한다.

검증할 곳

모델과 프롬프트 민감도를 기록하고, 중요한 판단은 실제 사람의
조사와 숙의로 확인한다.

- WP15는 인간 사전조사를 보완할 가능성을 제시한다. STEPI의 정책 증분효용을 입증한 자료는 아니다.

합성 페르소나로 미래를 묻는 법

STEPI 2045 프론티어 연구 | 방법, 실행, 결과와 한계



무엇을 만들었고, 무엇을 아직 말할 수 없는가

수집 완료와 타당화 완료는 다른 사건이다.

구축·실행

통계 앵커 카드, 190문항, 반복 실행, 모델 비교, 원자료
동결을 연결했다.

검증·공개

인간 타당화, 유형의 실재성, 정책 증분효용과 공식 결과 공개는
별도 게이트다.

- 내부 계산의 성공은 공개 승인이나 사람 대체의 근거가 아니다.



■ 읽는 순서

이 발표는 여덟 질문에 답한다

출발점에서 남은 검증까지, 하나의 증거 계보로 읽는다.

01 A 문제 왜 합성 페르소나인가	02 B 지형 NVIDIA, MIT, Stanford	03 C 설계 카드, 표본, 문항과 자극	04 D 실험 파일럿과 근거 깊이 비교
05 E 실행 본조사와 모델 분기	06 F 분석 반복성, 구조, 유형과 모델	07 G 인간 실제 조사와 워크숍 대조	08 H 활용 승인, 산출물과 다음 검증

■ 숫자의 분모와 지위를 먼저 읽고 해석한다.



세 가지 수를 혼동하지 않는다

레코드 수, 생성 런 수, 실제 사람 수는 교환할 수 없다.

단 위	현 재 연 구 의 예	해 석
합성 카드	본체 1,000개	고정된 조건 입력
생성 런	본체 3,000건	같은 카드의 3회 반복
모델 비교 카드	공통 200개	여러 모델에 반복 투입
인간 응답자	원본 1,303, usable 1,292	온라인 자발 참여자료

- 3,000런을 독립적인 3,000명으로 표기하지 않는다.



A

연구 문제와 의사결정

기술이 가능하다는 말과, 사람들이 받아들인다는 말 사이의 간격.

전문가의 미래상을 생활의 질문으로 바꾼다

또 하나의 가상 전문가가 아니라, 놓친 생활조건을 찾는 도구다.

01 미래상 어떤 기술과 제도가 실현되는가	02 생활조건 누구에게 혜택이 닿고 누가 배제되는가	03 수용조건 무엇이 보장되어야 받아들일 수 있는가	04 인간 확인 실제 조사와 숙의에서 무엇을 확인할 것인가
--	---	---	---



페르소나는 판정자가 아니라 관점을 탐색하는 렌즈다

사람을 대신 말하게 하지 않고, 사람에게 물을 질문을 만든다.

- 허용: 문항 사전점검, 반대논리 탐색, 시나리오 사각지대 발굴.
- 조건부: 워크숍 의제, 추가 조사와 표집 우선순위의 후보.
- 금지: 국민 찬성률, 미래 예측, 시민참여 대체, 단독 자원배분 근거.

-
- 운영 속도와 저비용은 타당도나 정책 효용의 대체 지표가 아니다.

설계는 한 번에 확정되지 않았다

사람 비교를 중심에 두면서 문항과 실행 구조가 바뀌었다.

06/19 01 생활세계·수요·사각지대 중심의 방법론 구상	07/02 02 재사용 킷과 고정 표본 구축	07/24 03 시계열 비교 우선에서 공통 코어+확장으로 전환	08/10 04 페르소나 190문항과 인간 비교층 확정	09/07 05 수집 완료 뒤 규칙·동결·공개 게이트 재정비
---	---	---	---	--

- 초기 제안은 현재 실행 사실과 구분한다. 원본 이력은 소급 수정하지 않았다.

7월 24일의 전환점 인간 문항이 비교의 중심이다

문항을 많이 생성하는 것보다, 같은 질문을 비교하는 것이 중요하다.

이전

과거 시계열 문항을 병치하고 100문항 파일럿으로 넓게 탐색.

이후

인간 공통 코어를 보존하고, 페르소나 전용 심층 문항을 별도 층으로 확장.

- 과거 문항은 폐기한 것이 아니라 활용 범위와 우선순위를 바꿨다.

세 산출물은 같은 질문을 반복하지 않는다

같은 증거를 쓰더라도 기관 해석, 연구 검증, 재사용 자산을 분리한다.

01

기관보고서

무엇을 구축했고 어떤 정책 질문을 제안하는가.

02

학술연구

이 생성계가 언제 맞고 어디에서 실패하는가.

03

재사용 자산

어떤 표본·문항·검증 절차를
다음 연구에 쓸 수 있는가.

- 출판과 외부 공개는 별도 권리 확인을 전제로 한다.

B

AI 페르소나의 지형

같은 페르소나라는 말 아래, 서로 다른 연구 대상과 평가가 있다.

데이터, 행위자, 개인모사 세 층을 나눠 읽는다

규모나 설득력 한 가지로 모든 접근의 성능을 비교할 수 없다.

접근	대표 사례	평가의 중심
합성 인구 카드	NVIDIA Nemotron Personas	통계 속성, 다양성, 사용권
상호작용 행위자	Stanford Generative Agents	기억·계획·반성, 행동의 개연성
실제 개인 모사	Stanford 인터뷰·설문 기반	보지 않은 응답의 재현
인구 동학	MIT AgentTorch / LPM	상호작용과 관측자료 보정
미래 자아 대화	MIT Future You	자기성찰과 심리적 효과

■ 각각의 성과가 곧 한국 정책 설문 타당성을 뜻하지 않는다.



NVIDIA의 700만은 700만 사람이 아니다

한국판 base는 100만 레코드 안에 7종 페르소나 서술을 담는다.

항 목	정 확 한 단 위	ST EPI에 서 의 의 미
1,000,000	합성 레코드	표본을 뽑는 원본 프레임
7,000,000	페르소나 서술 필드	직업·가족·취향 등 7종
26	레코드 필드	인구 속성과 서사 결합
CC BY 4.0	HF base 라이선스	출처표시 필수

■ NVIDIA / NAVER Cloud. 실존 인물의 개인정보 데이터셋이 아니다.



통계 앵커는 자유 생성의 쓸림을 제어한다

공공통계 → 구조적 표집 → 자연어 서술 → 검증의 순서다.

01 공공통계 인구·가구·지역의 기준	02 PGM 확률적 그래프 모델로 속성 구조화	03 LLM 서술 속성을 생활 맥락의 문장으로 확장	04 검증 형식, 분포, 다양성과 문화 오류 점검
---	--	---	--

■ 특정 변수의 독립성 가정과 누락된 상호작용은 공급자 카드에 명시돼 있다.



통계에 맞춘 카드도 한국 사회 전체는 아니다

관측 가능한 필드와 추정해 덧붙인 속성을 구분한다.

01 연령 19세 이상. 아동·청소년의 직접 관점은 빠져 있다.	02 성별 sex 이진 필드. 젠더 경험의 전 범위를 담지 않는다.	03 경제 base에 직접 소득·자산 필드가 없다.	04 교차성 직업과 여러 속성의 상호작용이 완전 복원된 것은 아니다.
--	--	---	---

■ housing_type은 주택 구조다. 월세·전세나 자산 수준으로 자동 번역하지 않는다.



Stanford 2023

그럴듯한 행동을 만드는 구조

25개 가상 행위자의 기억·반성·계획을 연결한 Generative Agents.

01

관찰

환경과 다른 행위자의 사건을 기록

02

기억

관련 경험을 선택적으로 검색

03

반성

경험을 상위 수준의 해석으로 압축

04

계획

기억과 해석을 다음 행동으로 연결

- 행동의 개연성 평가다. 실제 시민 설문에 대한 예측 정확도 검증과 다르다.

Stanford의 개인모사 실제 사람의 근거를 입력한다

1,052명의 실제 자료로 만든 에이전트와, 합성 인구 카드는 출발점이 다르다.

2024 v1

개인의 질적 인터뷰를 바탕으로 설문, 성격과 실험 반응을 모사했다.

2026 v3

인터뷰만, 설문만, 두 근거 결합을 분리해 비교한다. 동일 arXiv 기록의 개정판이다.

- 제목의 1,000과 실제 참가자 1,052를 구별한다. 미국 연구를 한국의 타당화로 인용하지 않는다.

'85% 정확도'라는 짧은 인용을 고친다

본인도 두 주 뒤 완전히 같은 답을 하지 않는다. 수치는 그 일관성으로 정규화됐다.

판본 · 조 건	보 고 수 치	분 모 의 의 미
2024 v1 인터뷰	85%	인간의 2주 재검사 일관성 대비
2026 v3 인터뷰	83%	보류 GSS 문항의 정규화 정확도
2026 v3 설문	82%	같은 비교 기준
2026 v3 결합	86%	같은 비교 기준
2026 v3 인구통계	74%	비교 baseline

■ 절대 정답률, 전인적 복제율, 한국의 정책 수용 예측률이 아니다.



MIT AgentTorch

사람 수보다 동학과 보정이 핵심

Large Population Models는 상호작용하는 대규모 인구의 변화를 모형화한다.

메커니즘

GPU 병렬 처리와 미분 가능한 시뮬레이션을 관측자료 보정에 결합한다.

STEPI와의 차이

독립적인 카드별 설문 응답 생성과, 네트워크를 통한 인구 동학 시뮬레이션은 다른 실험이다.

- 공급자의 계산 규모·속도 주장은 개인 선호 충실도의 독립 검증이 아니다. STEPI에 이 엔진을 적용한 것은 아니다.

MIT Future You

미래를 예측하는 쌍둥이가 아니다

미래 자아와의 대화는 자기성찰을 돕는 개입이다.

- 사용자의 정보를 바탕으로 미래 자아의 서사와 대화 경험을 만든다.
- 평가의 대상은 불안, 정서와 미래 자아 연속성 같은 심리적 변화다.
- 미래에 실제로 어떤 삶을 살지 맞힌 예측 실험으로 읽지 않는다.

다른 접근은 비교 대상이지 채택 실적이 아니다

벤치마크 후보, 제품 소개, 실제 실행을 구분한다.

계 열	얻 는 관 점	이 연구에서의 지 위
Silicon sampling	조건부로 인간 패턴을 모사할 가능성	문헌 배경
ManyPerson	한국 통계 기반 여론 시뮬레이션 사례	초기 비교 후보, 실행 근거 없음
MatrAlx	카드×모델×과제×검증기 결합	위키 기반 구조 비교
No-persona baseline	페르소나 서사의 증분효과 식별	본조사 후속 과제

■ MatrAlx의 역할준수는 사람 예측 타당도가 아니다. MIT 코드 라이선스는 MIT 대학 소속을 뜻하지 않는다.



선행연구가 주는 공통 설계 원칙

가능성과 실패를 함께 읽어야 검증 설계가 나온다.

01

근거

인구통계, 서사, 인터뷰를 구분한다.

02

도구

모델과 프롬프트를 실험
조건으로 기록한다.

03

비교

평균뿐 아니라 분산,
개인차와 하위집단을 본다.

04

한계

내부 일관성을 외적
타당도로 바꾸지 않는다.

- '잘 모사한다'는 주장은 언제, 누구를, 무엇에 대해의 범위를 포함해야 한다.

C

표본과 조사도구

프롬프트 한 장이 아니라, 표본·자극·문항·검증 계약을 만든다.

표본은 실행 전에 고정한다

원본 프레임 → 할당 셀 → 고정 명부 → 모든 반복에 동일 투입.

01

원본

Nemotron-Personas-Korea base

02

셀

age_bin × sex × province
비례할당

03

명부

쿼터 1,000, seed 42

04

반복

같은 카드에 r1, r2, r3 적용

- seed는 표본 추출의 재현성을 고정한다. LLM 재호출의 동일 출력을 보증하지 않는다.

쿼터 1,000과 레드팀 20 서로 다른 일을 한다

넓은 반응 탐색과 극단 사각지대 탐색의 분모를 합치지 않는다.

쿼터 1,000

합성 프레임의 인구 속성 구성을 넓게 배치한다. 확률표집된 시민 1,000명이 아니다.

레드팀 20

극단·취약 조건을 목적으로 고른다. 본체와 교집합이 없는 별도 명부다.

- 레드팀의 높은 우려 비율을 전체 합성 프레임이나 국민에게 일반화하지 않는다.

'분포 편차'에도 기준 집합이 필요하다

원본 프레임과의 정합은 실제 국민 태도와의 정합이 아니다.

점 검	확 인 내 용	주 의
원본 무결성	100만 행, 26필드, 고유 식별자	형식·도메인 점검
표본 재현	seed 42 재추출 1,000/1,000	동일 표집 코드 조건
P2 요약	최대 마지널 편차 1.93%p	해당 감사 변수 범위
직업 부표	무직 편차 2.33%p	1.93을 모든 변수 최대로 일반화 금지

■ 후대의 '쿼터 편차 0'과 원본 프레임 마지널 편차는 정의가 다르다.



잘못된 원본 설명을 실제 필드로 교정했다

풍부한 서사가 있어도 존재하지 않는 필드를 만들지 않는다.

초기 표현	확인된 사실	적용
온라인쇼핑판매원 19.8%	정본 라벨 약 0.50%	초기 수치 폐기
성씨 분포	base에 성씨 필드 없음	서사 가명과 필드 구별
주택형태=점유형태	아파트 등 구조형	자가·전세 추정 금지
학력 47.5+29.6~82	산술합은 77.1	오기 그대로 복제 금지

■ 원자료 검증 PASS는 모든 산문 해석의 PASS가 아니다.



조사 대상은 여섯 미래상이다

아래 명칭은 연구의 자극 체계다. 정부 최종 확정 비전이라고 부르지 않는다.

01 S1 지능국가 AI·양자 기반 국가혁신	02 S2 초장수 생명안전 예방의료와 재난 대응	03 S3 에너지 선도 자립, 무탄소, 자원순환
04 S4 생활·경제권 확장 공간 제약을 넘는 이동과 활동	05 S5 롤메이커 기술 질서와 규칙 설계	06 S6 포용·상생 혜택의 접근과 배분

■ S0는 기준 자극이다. 여섯 평가 미래상에 더해 일곱 정책안으로 세지 않는다.



미래상을 같은 크기의 상자로 보지 않는다

인프라, 개별 영역, 규칙과 포용이라는 층위가 함께 들어 있다.

- 본조사에서 S6는 개별 평가 대상이다.
- 청년워크숍 사전설문에서는 5개 미래상과 포용·배제의 공통 관점으로 다뤘다.
- 워크숍 단문 요약 노출을 본조사 장문 자극 노출과 동일시하지 않는다.

-
- 5도메인과 6자극문의 대응은 문서와 노출 이력을 확인한 뒤 확정한다.

190문항은 두 층으로 구성된다

공통 코어 73 + 페르소나 확장 117 = 190.

층	주요 블록	수
STEPI 코어	OPEN, A~G	73
CS 확장	CS-A~F, Q, T, An, M, ST, O	117
합계	고유 문항 id	190
6축 심층	CS-M	55

■ 여기의 L1/L2는 문항 층이다. 분석 단계 L0~L3의 L1/L2와 다른 표기다.



사람과 비교되는 공통 문구는 OPEN7이다

기대 국가상, 미래 과제, 과학기술 역할, 개인 기대, 우려, 현재 경쟁력, 미래 도전.

01

문구

OPEN1~7 = 사람 P2B1~7 원문

02

순서

미래상 보기 전 응답하도록 배치

03

단위

문항별·응답자별 개방형 텍스트

04

비교

같은 코드북으로 주제 언급을 대조

- 인간 자료에는 본조사 6축 리커트가 없다. 공통 문구가 전 척도 비교를 열지 않는다.

네 가지를 따로 묻고 수용성 여섯 축을 심화한다

바람직함, 실현 가능성, 개인 영향과 시급함을 한 점수로 압축하지 않는다.

B 블록

미래상마다 바람직성·실현가능성·개인영향·시급성의 4문항.

CS-M

기대 12, 체감성 6, 소외 6, 불안 8, 신뢰 7, 수용도 16문항의 6축.

- 6축 문항 수 합은 55다. 모든 축×미래 점수가 같은 문항 수를 갖지는 않는다.

측정하려는 전제를 먼저 정답으로 주입하지 않는다

'AI 최종책임은 인간에게'라는 전제 자체가 측정 대상이었다.

- 국가 체제, 지구 활동 공간, 물리법칙 전제만 시스템에 주입했다.
- 인간 최종책임 전제는 주입하지 않고 관련 문항으로 물었다.
- 정책 명제에 대가 절이 포함되면 동의율의 절대치 해석을 제한한다.

-
- 프롬프트로 선행 합의를 주입한 뒤 같은 합의를 발견했다고 말하지 않는다.

순서를 바꾸되 무엇이 섞이는지 기록한다

무작위화가 있는 부분과 없는 부분을 구분한다.

장 치	구 현	식 별 경 계
OPEN7 선행	자극 노출 전 배치	전역 문맥 누출은 별도 감사
E 보기	카드별 셔플, 반복 내 고정	canonical 번호로 저장
C2	C1 선택 제외	교차문항 제약 검증
CS-M 순서	라운드별 로테이션	순서×시점×라운드 교락

■ 순수 순서효과가 아니라 round-order bundle, 라운드·순서 결합조건의 차이다.

D 파일럿에서 배운 것

작은 실험의 관찰을 큰 타당성 주장으로 늘리지 않는다.

파일럿 v0.3

50개 카드로 조사 흐름을 점검했다

50개 합성 카드 × 100문항 × 1회 응답.

항 목	내 용
전신	07/05 v0.1 모빌리티 28문항에서 확장
자극	S0 기준문과 S1~S6 미래상
결과	구조 검증, 반응 지형, 시연용 볼트
분석	수용도, 입장, 양극 문항, 규칙 기반 유형
한계	단일 회차, 본조사와 문항 버전 다름

■ 50×100×7을 전체 응답 수로 곱하지 않는다. 모든 문항을 자극마다 반복한 설계가 아니다.



파일럿의 우열은 인간 수용률이 아니다

M블록 수용도 평균에서 관찰된 합성 지형이다.

미 래 상	평 균 (1 ~ 5)	읽 는 법
S6 포용·상생	4.12	이 파일럿 생성계 안의 상대값
S3 에너지	3.47	같은 문항 버전과 표본 조건
S2 초장수	3.20	사람 기준 검증 아님
S5 룰메이커	3.18	정책 우선순위 확정 아님
S4 생활권	3.05	본조사 수치로 이월 금지
S1 지능강국	2.53	국민 거부율 아님

■ 이 표는 과거 실험의 기록이다. 본조사 승인 결과로 대체하지 않는다.



파일럿 6유형은 가설이었다

컷포인트로 분류한 이름과 데이터에서 확인된 실재 유형은 다르다.

01 구조 소외형 14 / 50	02 적극 수용형 13 / 50	03 전환 불안형 8 / 50
04 조건부 실용형 7 / 50	05 유보·경계형 5 / 50	06 기대-체감 갭형 3 / 50

- 이 여섯 이름을 본조사 분류에 그대로 붙이지 않는다. 구성 변수로 다시 집단차 검정하지 않는다.



'REAL'이라고 부른 응답도 사람의 실답은 아니었다

두 종류의 근거를 입력한 LLM 응답끼리 비교한 실험이다.

문서 근거

한 사람의 확인된 문서와 지론을 제공한 생성 응답.

통계 클론

인구 속성과 직업 활동이 유사한 합성 카드를 제공한 생성 응답.

- 실제 당사자의 직접 설문 응답은 없었다. 인간 정확도 실험으로 소개하지 않는다.

비슷한 평균과 같은 판단은 다르다

v0.3 비교에서 척도 거리는 작아도 가치 선택은 같았다.

비교 지표	관찰	의미
Likert 83문항	MAE 0.51	평균절대오차
±1 이내	80 / 83	정도의 근접성
양극 동일극	11 / 17	가치 방향의 불일치 잔존
미래상 입장 일치	2 / 6	집계 근접≠개인 판단 복제

■ 한 사례의 내부 비교이며 ‘사람을 맞혔다’는 결론을 지지하지 않는다.



근거 깊이를 다섯 조건으로 바꿨다

문서 A, 프로필 B, 통계 클론 C, 무작위 D1·D2.

01

A 문서

비교 기준으로 사용한 생성 응답

02

B 카드

같은 인물의 프로필 요약

03

C 클론

통계적 최근접 합성 인물

04

D 무작위

서로 다른 무작위 카드 두 개

- 5조건×190문항, 950개 문항 응답. 각 조건 1회라 반복 불확실성을 추정하지 못한다.

카드 근거는 문서 근거에 더 가까웠다

척도 148문항에서 A 대비 거리와 일치도를 비교했다.

조 건	M A E	완 전 일 치	± 1 이 내
B 프로파일	0.19	81%	100%
C 통계 클론	0.49	55%	96%
D1 무작위	0.61	48%	91%
D2 무작위	0.96	36%	78%

- 방향성 가설을 얻었다. 근거 깊이 외 인물 속성도 달라, 순수한 인과효과로 단정하지 않는다.



작은 실험이 입증하지 않은 것

‘프로파일이 집계 타당성을 입증했다’는 초기 해석은 과도하다.

- 기준 A도 문서 기반 생성물이라 인간의 정답이 아니다.
- 무작위 조건은 인구 속성까지 달라진다. 근거 깊이만의 통제가 아니다.
- 단일 인물·모델·회차에서 얻은 일치도를 전체 1,000개 카드로 일반화하지 않는다.

-
- 남은 성과는 후속 baseline과 실제 인간 응답이 필요하다는 구체적 설계 근거다.

E 본조사 실행과 동결

응답을 만드는 모델과, 실험을 통제하는 코드를 분리한다.

코드가 통제하고 모델은 응답값을 만든다

생성의 자유와 실험의 불변조건을 서로 다른 책임으로 둔다.

01

입력 계약

카드·문항·자극·순열 고정

02

생성

LLM이 조건부 응답값 출력

03

검증

id·범위·선택 수·흐름 검사

04

저장

채택 원본·로그·provenance 보존

- 검증기는 의미의 진실을 판정하지 않는다. 구조 검증과 내용 검증은 별개다.

구조 검증은 엄격하지만 충분하지 않다

빠짐없이 형식에 맞는 답과, 타당한 답은 다르다.

기계 점검	기대 조건	보증하지 않는 것
문항 id	190개, 누락·중복 없음	사람과의 일치
값	척도 타입과 범위	가치 판단의 진실
선택	$E1=5, E2=3, E3\leq3$	정책 우선순위 타당성
관계	$C2\neq C1$, 제시 순서 일치	문화적 충실도

- 실패 재시도와 복구 이력도 남겨야 최종 성공률의 선택 과정을 볼 수 있다.



4,650과 60 같은 합계로 덮지 않는다

본체와 비교 암의 동결 뒤, 레드팀은 별도 증분으로 남았다.

묶음	산식	채택 응답
본체	1,000×3회	3,000
통제	동일순서 통제50	50
모델 암	200×8조건	1,600
기본 동결	위 세 묶음	4,650
레드팀 증분	20×3회	60

- 기본 동결 4,650에 레드팀 60은 포함되지 않는다. 두 묶음을 함께 세면 4,710건이다.



첫 라운드는 단일 모델이 아니었다

요청한 이름이 아니라 실제 응답의 계보로 분석세트를 나눈다.

라운드	관측 모델	건수
r1	Opus 5	757
r1	Fable 5.1 / Fable 5	214 / 29
r2	Opus 5 [1m] / Opus 5	782 / 218
r3	Opus 5 [1m]	1,000

■ 위 모델 표기는 이 연구의 실행 로그 라벨이다. 과거 r1의 미관측 티어를 소급 확정하지 않는다.



실행 중 전환도 측정도구의 변화다

조건 전환을 숨기면 반복성의 분모가 틀어진다.

08/31 01 본조사 시작, r1 모델 기본값 혼입	09/03 02 모델 명시, 컨텍스트 티어와 실패 경로 점검	09/04 03 라운드별 문항 파일 모드와 [1m] 명시	09/05~06 04 r2 티어34건 재실행 교체, 재정렬32건 포함	09/07 05 본체·모델 암 동결, 레드팀 별도 완료
--	--	--	---	---

- 낮은 r1 일치로 재실행한227건은 r1 Fable 혼입 교락이 밝혀져 교체 보류·보존했다. 권한모드를 품질 저하 원인으로 단정하지 않는다.



반복 분석의 정본은 C1 757개 카드다

같은 모델 계열과, 완전히 같은 실행조건은 다르다.

세트	고유 카드	역할
r3 전체	1,000	가장 균질한 단면
C1	757×3회	r1 Opus와 후속 반복의 교집합
r2↔r3	1,000×2회	두 라운드 민감도
실행조건 동질	374×2회	확인된 실행조건 부분집합
r1 Fable	243	별도 표, C1에 합산 금지

- C1도 티어·시점·입력 형식이 완전 동일한 실험은 아니다.



control50은 완벽한 동일조건 통제가 아니다

같은 순서로 다시 물었어도, 실행 티어가 달랐다.

- 200K 티어 경로의 50개 응답으로 남아 있다.
- r1과 모델이 맞는 쌍은 일부이며 실행시점도 달랐다.
- control50은 티어 민감도, 재정렬32쌍은 하위집합 비교다. 어느 쪽도 전체 반복의 천장이 아니다.

-
- 추가50런 재실행은 보류다. 32쌍의 paired provenance-구간 해석은 별도 감사 대상이며 이 발표 제작에서 연구를 재실행하지 않았다.

모델 암은 다수결용 여덟 여론이 아니다

동일 카드를 다른 생성 도구에 넣어 도구 민감도를 본다.

구 성	역 할
Opus [1m] 인라인	본체 파일모드와의 형식 비교 기준
Sonnet, Gemini 2종	다른 계열·크기 조건
GPT Sol / Terra	서로 다른 실행 모델 조건
Grok 2종	다른 모델 조건
main r3 동일200	기존 본체의 대응 행, 추가 수집 아님

- 1,800 분석행의 독립 단위는 공통 카드 200개다. 인라인과 파일모드는 프롬프트 해시가 다르다.



동결은 원본과 파생물을 나누는 일이다

원본을 다시 쓰지 않아야 해석이 바뀌어도 당시 증거가 남는다.

01

원본

채택 응답, 입력과 실행로그

02

명부

파일별 해시와 분석세트

03

파생

새 분석코드와 산출물 버전

04

공개

허가된 지표만 릴리스

- Git revision만으로 dirty 작업내용을 재현할 수 없다. 원본 해시와 당시 코드 스냅샷이 필요하다.

F 실제 분석과 검증 경계

다음 장들은 공식 결과 릴리스가 아니라, 이미 존재하는 탐색 진단의 검토다.



분석은 아홉 기능으로 질문을 분해한다

자료 준비 뒤, 반복·구조·집단·관계·유형·예측·모델·사람을 따로 점검한다.

모듈	질문	도구
준비 / M2	자료가 맞고 반복되는가	키·코호트, ICC·일치율
M3 / M4	축이 분리되고 집단차가 남는가	EFA·CFA, 효과크기·조정모형
M5 / M6	관계·유형이 식별되는가	반응표면, 군집·LCA
M7 / M8	예측·모델 변경에 무엇이 남는가	교차검증, 짝지은 모델 대비
M9	사람 자료와 무엇이 다른가	공통층 주제 비교·민감도

- 모듈 파일 존재, 계산 완료, 검토 완료, 공식 공개는 네 개의 서로 다른 상태다.



최신 ICC는 예전 근사값을 대체해서 읽는다

C1 757개×3회, 절대일치 ICC(2,1), UUID 묶음 bootstrap 5,000회.

측	ICC (2 , 1)	95% 구간	참조 밴드
기대	0.531	[0.481, 0.577]	조건부
체감성	0.735	[0.705, 0.762]	조건부
소외	0.817	[0.792, 0.839]	안정
불안	0.520	[0.470, 0.563]	조건부
신뢰	0.309	[0.251, 0.365]	불안정
수용도	0.528	[0.490, 0.565]	조건부

■ 95% percentile 구간. 수치는 탐색 점검값이며 T01 서명·T20 정식 판정의 완료를 뜻하지 않는다.



점추정 밴드와 정식 판정은 별개다

≥0.75, 0.50~0.75, <0.50의 기존 참조선을 사후에 바꾸지 않는다.

01 안정 참조 소외 1축.	02 조건부 참조 기대, 체감성, 불안, 수용도 4축.	03 불안정 참조 신뢰 1축.
--	---	---

■ 기대·불안·수용도 CI는 0.50, 체감성 CI는 0.75를 지난다. 구간을 새 판정 규칙으로 쓰지 않는다.



높은 정확일치와 낮은 ICC는 양립한다

모두 비슷한 답을 하면 정확일치는 높아도 개인차를 안정적으로 구분하지 못한다.

정확일치율

같은 문항에서 같은 값을 반복했는가.

ICC

개인 간 차이 대비 반복 오차와 조건 차이가 얼마나 작은가.

■ ICC(2,3)의 평균화 효과를 ICC(2,1)의 실패를 구제하는 데 사용하지 않는다.



r3의 반응은 좁은 범위에 압축됐다

개인별 축 평균의 평균과 표준편차다. 문항 자체의 분산과 다르다.

축	평균 (1 ~ 5)	개 인 평 균 S D	고 유 값 수
기대	3.988	0.137	13
체감성	2.855	0.379	12
소외	3.942	0.350	15
불안	3.948	0.159	10
신뢰	2.144	0.150	9
수용도	3.508	0.162	17

■ 저분산 축을 z표준화하면 작은 차이가 과장될 수 있다. 합성 생성계의 분포이지 인간 모집단 분포가 아니다.



M3 구조 분석

여섯 축이라고 여섯 요인은 아니다

설계 이름의 개수와 경험적 측정구조를 따로 검증한다.

01 분할 r3 개발500 / 확인500	02 탐색 평행분석, minres EFA, oblimin	03 확인 사전 지정 CFA, 해의 적격성 검사	04 제한 희소 범주, 분산 압축, 순서형 추정의 불안정
---	--	---	--

- 인간에게 같은 척도를 수집하지 않아 인간-AI 측정동일성은 검정할 수 없다.



부적격 CFA 해는 성공으로 수선하지 않는다

확인 반의 6요인 모형에서 부적격 해가 기록됐다.

- 잠재공분산 비양정치, 절대 요인상관 1 초과 등은 해석 중단 사유다.
- 요인 병합의 경계 귀무가설에 통상의 카이제곱 차이검정을 쓰지 않는다.
- 탐색적 재명세와 확인 표본에서의 검증을 분리한다.

-
- 09/10 marker 경계조건 수정은 모의자료로 검증됐다. 실제자료 재실행 결과가 아니다.

M4 집단차 유의성보다 크기와 일관성

단순 차이, 조정 차이, 반복 조건의 일관성을 분리한다.

01 이변량 Welch, 효과크기, 원척도 차이	02 조정 OLS-HC3, 조정평균, partial R²	03 반복 UUID 임의효과와 round 고정효과	04 판정 축 신뢰도, 동일 대비 방향, 다중검정 가족
---	---	--	---

■ 작은 원척도 차이가 저분산으로 큰 표준화 효과처럼 보일 수 있다. 임시 SESOI는 승인 기준이 아니다.



M5 반응표면 곡면이 있다고 인과가 생기지 않는다

RSA는 두 입력의 일치·불일치와 곡률을 보는 다항회귀다.

페르소나 내부

기대×신뢰, 소외×체감성 등과 수용도의 관계를 모델링한다.

사람×AI 결합

층별 언급률은 소수 층의 집계변수다. 1,000행으로 펼쳐도 독립 정보가 늘지 않는다.

- 같은 시점의 생성 응답으로 매개·조절을 적합해도 인과 메커니즘으로 해석하지 않는다.

통계 기준을 넘어도 정식 검증을 못 넘을 수 있다

M5 내부 후보 기준과, 사용 점수의 안정성 기준이 따로 작동했다.

- 모듈 내부 통계 후보 7건, 안정성 게이트를 거친 본문 승격 0건.
- 조건부·불안정 점수를 사용한 모형을 확증 결과로 올리지 않는다.
- 진단 모듈의 승격 플래그 자체도 프로젝트의 사람 서명을 대신하지 않는다.

-
- 기존 출력은 탐색 산출물이다. 정책효과나 인간 행동의 인과관계로 일반화하지 않는다.

M6 유형화 잘 나누는 것과 다시 나누는 것

군집 내부 재표집 안정성과 라운드 간 재현성이 다르게 나타났다.

검 사	관 찰 값	범 위
k=2 silhouette	0.339	r3 분리 정도
k=6 silhouette	0.213	강제 6군집
k=2 재표집 ARI	0.961	같은 r3 재표집
C1 r2 / r1 ARI	0.345 / 0.357	r3 중심으로 재배정
M-arm ARI 중앙값	0.086.	공통 200카드, 모델간

■ ARI는 분류 일치 지수다. 재표집 구간은 국민 유형의 신뢰구간이 아니다. 이산 유형의 존재는 미확인이다.



같은 여섯 이름을 사람과 AI에 붙일 수 없다

공통 유형의 출처 간 비교가능성이 확보되지 않았다.

- k=6 강제 적합은 파일럿 유형의 직접 재현 검증이 아니다.
- 사람·합성 태그의 잠재계층은 길이와 출처 분리에 민감했다.
- 내용적 유형명은 구조와 출처 비교가능성 확인 뒤에 붙인다.

M7 기계학습 생성계의 패턴을 예측한다

같은 사람·카드의 문항이 학습과 평가에 섞이지 않도록 묶는다.

01

수용도 예측

LightGBM과 ridge를 비교한다.

02

출처 판별

인간과 합성의 OPEN7
텍스트를 분류한다.

03

라운드 판별

r1과 r3의 문체 변화를
기준선으로 본다.

04

검증 단위

UUID·응답자 단위 5-fold,
변환은 fold 안에서 학습.

- 합성 응답 예측 성능이 실제 국민 행동 예측 성능은 아니다.

인간과 합성 텍스트는 쉽게 구별됐다

내용뿐 아니라 길이만으로도 출처 구별력이 높았다.

특징	AUROC	검증 단위
길이만	0.9768	응답자 묶음 5-fold
내용+길이	0.9999	응답자 묶음 5-fold

■ M7 해당 경로는 사람576+합성855 응답자다. M9 사람573과 결측 성별 처리 분모가 달라 같은 표본이라고 쓰지 않는다.



M8 모델 민감도 평균순위와 개인 응답을 분리한다

미래상의 집계 순위가 비슷해도, 개별 응답값은 크게 다를 수 있다.

집계 층

수용도 평균의 순위, 미래상 쌍의 부호 보존.

개인 층

동일 카드의 정확일치, MAE, 군집 배정, 축별 이동.

- 모델 간 일치율 차이를 분산성분비 '5~10배'로 바꾸지 않는다.

같은 카드라도 모델이 바뀌면 답이 달라진다

인라인 Opus 대비, 카드별 정확일치율의 평균이다.

비교 조건	정확일치율	95% UUID 구간
Sonnet	0.599	[0.591, 0.607]
Gemini Pro	0.590	[0.582, 0.598]
Gemini Flash	0.657	[0.650, 0.665]
GPT Sol	0.557	[0.550, 0.565]
GPT Terra	0.592	[0.583, 0.601]
Grok reasoning	0.496	[0.487, 0.504]
Grok 4.6	0.644	[0.636, 0.652]
본체 Opus 파일모드	0.812	[0.805, 0.819]

■ 본체 동일200의 파일모드 비교를 함께 본다. 확률표집 오차나 사람 정확도의 표가 아니다.



미래상 순위에서 보존된 것과 바뀐 것

S6 상단, S3 다음, S1 하단은 보존됐고 중간 순위는 달랐다.

- 8모델×6미래상의 Kendall W = 0.941 (순위 일치 진단).
- 6개 미래상만의 순위상관은 거칠고 이산적이다.
- 동일 자극의 규범적 문구가 여러 모델의 공통 반응을 만들 가능성도 남는다.

-
- 공통 순위를 국민 선호나 정책의 정당한 우선순위로 발표하지 않는다.

9월 10일에 고친 것은 기존 결과가 아니라 경계조건이다

실제자료 M3~M9 재실행은 하지 않았다.

모 들	좁 은 수정	검 증
M3	marker 식별 불가·모형 밖 값 차단	M3/M6 30개 테스트
M6	빈 군집 포함 최소 비중 집계	동일 검증 묶음
M8	누락 p 가족 크기, 분산0 미정의 처리	M8/M9 65개 테스트
M9	빈 층을 0 대신 NaN 전파	동일 검증 묶음

- M4 집합 정합, 일부 LCA·렌더·GEE 등은 미검토·보류 범위다. 전체 감사 완료가 아니다.



G

실제 사람 자료와 대조

사람 자료가 있다는 사실만으로 모든 비교가 가능해지지 않는 다 .

1,292명은 국민 대표 표본이 아니다

원본 1,303 중 기존 필터의 usable은 1,292다.

특 성	확 인	합 의
모집	온라인 자발 참여	자기선택 편향
연령	30대·40대 비중 큼	합성 프레임과 구성 차이
문항	P2B1~7 개방형	OPEN7 제한 대응
척도	동일 6축 없음	인간-AI 축 분산비 미식별

- 사람 자료의 한계를 인정하는 것이 합성 응답을 사람과 동등한 관측으로 승격시키지는 않는다.



같은 필터를 적용해도 같은 영향을 주지 않는다

최소 응답 길이와 반복문자 규칙은 사람을 훨씬 더 많이 제외했다.

집 단	선택 단 계	남 은 응 답 자
사람	원본	1303
사람	기존 usable	1292
사람	공통 길이·반복문자 통과	726
사람	공통 6층	573
합성 r3	공통 규칙 통과	1000
합성 r3	공통 6층	855

■ M9 공통 6층은 30대·40대·50+ × 남녀다. 필터를 실제 성실도의 정답으로 취급하지 않는다.



M9는 공통층에서 주제 언급 차이를 본다

인구 구성을 맞추는 것과 문체·코딩 차이를 제거하는 것은 다르다.

01 대응 같은 개방7 문구	02 코딩 18주제 규칙사전, tag_any와 문항별	03 표준화 공통 6층 동일가중의 합성-인간 차이	04 민감도 필터, 연령, 가중, 길이, 라운드
--	--	--	---

■ D06 사용자 비교 승인과 STEPI 승인·위임 기록은 별개다. 본 발표의 공개와 원연구의 정식 비교 판정은 별개다.



주제 차이는 곧 관심 차이가 아니다

생활어와 정책어, 응답 길이와 사전 누락이 함께 움직인다.

- 동일 키워드 사전은 재현성을 높이지만 내용 타당성을 보증하지 않는다.
- 층화와 가중은 선택·문체·측정오류를 모두 제거하지 못한다.
- 출처를 가린 이중코딩과 불일치 조정이 후속 과제다.

-
- 차이의 관찰과 원인의 확정을 분리한다. 언급하지 않았다고 그 주제에 무관심한 것도 아니다.

'2020 시계열'의 원자료는 2019 조사였다

보고서 연도와 조사 연도, 문구와 모집 방식을 함께 확인한다.

과거

2019년 조사. 일반인, 산업계, 전문가의 질문과 표본이 다르다.

현재

2026년 자발 참여자료. 기대·도전·우려의 문구가 과거와 완전 동일하지 않다.

- 2019→2026은 7년 간격이다. 부분 r1 594건을 사용한 옛 표를 최신 본조사 표로 복사하지 않는다.

워크숍 수령은 두 레인으로 나뉜다

개인 설문과 공동 산출물은 분석 단위와 연결 가능성이 다르다.

레인	자료	현재 지 위
개인자료	사전·사후 설문, 현장 투표 등	수령 기록, 연결·동의 범위 검수
공동자료	월드카페 XLSX	구조 추출 완료, 코딩 대기
공동자료	2045 청년의 하루 XLSX	이야기·장면 추출, 코딩 대기

- 공동 파일의 개인키 부재를 모든 개인자료의 영구 연결 불가로 일반화하지 않는다.



161개 진술과 30장면은 191명의 응답이 아니다

행, 진술, 이야기, 장면의 중첩 구조를 그대로 보존했다.

산 출	수 량	지 위
월드카페	진술 후보 161	포함 여부 hold
R4 행	5	예시 여부 unknown
청년의 하루	이야기 10	공동작성 단위
장면	30	이야기 안에 중첩
구조 테스트	24/24	주제 타당화 검사 아님

- 모든 레코드 coding_not_performed. 시트→S축 매핑은 null로 남겼다.



숙의 자료는 독립 설문처럼 셀 수 없다

자율선택, 상호 노출, 공동 작성과 앞선 활동의 영향이 있다.

- 월드카페의 3턴은 참가자 간 상호작용을 포함한다.
- 같은 이야기의 3장면은 독립 관측이 아니다.
- 두 워크북의 일치도 순차 의존을 가진다. 독립 확증 두 개가 아니다.

-
- 제안 코드북, 독립 2패스 코딩, 사람 조정의 절차는 아직 실행 결과가 아니다.

H 무엇을 남기고, 무엇을 더 검증할까

계산 결과의 양보다, 주장과 증거의 연결을 완결한다.

결정 원장은 권고와 승인을 나눈다

승인 범위를 좁게 읽어야 연구의 경계가 유지된다.

결정	확인된 상태	남은 것
D01 / D17	축별 규칙·분석세트 승인	T01 서명, T27 연결
D02	추가 통제 실행 보류	분석 뒤 재상정
D03	추가규칙 등록 방식 승인	세부값·실행 일괄승인 아님
D05	레드팀 3회 완료	원탁회의 미착수
D06	사용자 승인	STEPI 승인·위임 대기

■ 09/09 ICC5000의 내부 실행 예외는 정식 판정과 외부 공개 허가가 아니다.



L0에서 L3까지 앞 단계가 뒤 단계를 보증하지 않는다

자료 건전성 → 생성 반복성 → 비교가능성 → 인간 기준과 정책 효용.

01 L0 건전성 채택 집합, 해시, strict 검증	02 L2 도구 검증 반복.순서.모델.내용 감사	03 L1 반응 분석 허용된 점수와 유형만 해석	04 L3 외부 대조 사람.워크숍.과거자료의 범위별 비교
---	---	---	--

■ 분석 순서는 번호순이 아니라 L0→L2→L1→L3다. 조건 미충족은 미실시·미식별로 남긴다.



후속 검증은 증분효과와 실패 조건을 겨냥한다

기존 데이터를 본 뒤 설계한 검증을 사전등록 결과라고 부르지 않는다.

01 Baseline no-persona / 인구통계만 / full-persona	02 순서 균형 같은 라운드 안 split- ballot 또는 Latin square	03 인간 기준 같은 시점·자극·문항의 독립 holdout	04 정책 효용 후보 질문의 신규성·타당성 ·실행가능성 맹검 평가
--	--	---	---

- 새 확인적 연구는 규칙·권리·예산 승인 후 별도 wave로 설계한다.



원탁회의와 backcasting 생성물에서 정책근거까지의 간격

사각지대 후보를 곧바로 공식 과제로 승격하지 않는다.

01

대립 구성

이해상충 관점으로 쟁점 후보 생성

02

증거 대조

원자료·문헌·실제 참여 경험 확인

03

역산

2030·2035·2040 준비과제 후보

04

사람 검토

채택·수정·기각 이유 기록

- 레드팀 응답60 수집과 정책 분석 완료는 다르다. 원탁회의 생성은 확인된 최신 기록상 미착수다.

납품 준비는 연구 결과의 승격과 별개다

볼트, 보고서, 발표덱이 같은 허가된 지표와 근거를 가리켜야 한다.

산출물	연결 원칙	현재 주의
보고서	허가된 metrics 릴리스	공식 결과 릴리스 미완
페르소나 볼트	카드↔문항↔증거	시연 존재≠결과 승인
발표덱	동일 주장·분모·상태	공개 발표본, 분석 검증 중
HWPX / PDF	형식 변환·시각 검수	기관 양식·인수 별도

■ 데이터셋 라이선스가 클라이언트 자료·참가자 발언·최종 성과물의 공개 권한까지 만들지는 않는다.



이 연구가 지금 확실히 남긴 것

성공 순위 하나보다, 생성계를 검증 가능한 대상으로 만든 구조가 남는다.

01	02	03	04
조사 자산	실행 증거	실패 지도	다음 질문
고정 카드와 기계가독 문항·자극	모델·조건·실패·복구의 계보	반복·구조·유형·사람 비교의 한계	사람에게 확인할 의제와 후속 검증

■ 세계 최초, 국민 대표성, 전인적 복제, 정책효과를 입증했다고 주장하지 않는다.



출처를 따라가면 주장도 되돌릴 수 있어야 한다

외부 연구, 역사 실험, 최신 실행과 내부 진단을 분리했다.

층	주요 근거
외부 원문	NVIDIA 카드, Stanford 2023·2024v1·2026v3, MIT LPM·Future You
역사 실험	파일럿47, 트윈42, gradient60
설계·실행	문항53, 실행68, 동결 명부, 계획85
최신 검토	ICC5000 09/09, 워크숍93, 경계조건94
숫자 주입	원본 JSON 또는 보존된 진단 표의 지정 행

■ 상세 파일·해시·위치·정정사항은 마스터와 근거원장에 있다. 원자료·개인정보는 데크에 담지 않았다.



사람을 대신하지 않고 사람에게 묻는 일을 개선한다

합성 페르소나의 가치는 답의 양이 아니라, 더 나은 질문과 검증 가능한 한계다.

01 다음 T01 서명과 세부규칙 잠금	02 다음 사람 비교 요건·코딩·권리 확인	03 다음 정식 지표 검증 후 발표 내용 갱신
--	--	--



구축한 것은 검증의 기반 남은 것은 사람과 효용의 증거다

반복 가능한 생성계, 실패조건 지도, 후속 질문의 후보를 다음 검증으로 연결한다.

01

확인된 자산

고정 카드·문항·실행 계보와 반복
·모델·구조 진단을 연결했다.

02

아직 미확인

인간 대표성, 공통 유형의 실재성, 정책
판단의 증분효용은 입증되지 않았다.

03

다음 행동

사람 기준의 독립 비교, 문항 사전점검,
실무자 맹검 평가로 용도별 효용을 검증한다.

- 자료 공개는 검증 완료와 다르다. 미충족 조건에서는 사람 재현·정책효과 주장을 줄인다.

기업교육과 코칭

기업의 질문을 AI 활용의 출발점으로 바꿉니다.

POSITIONING

도구와 교육을 실제 운영으로 연결합니다

AI 도입의 병목은 기능 부족보다 조직의 맥락·지식·책임이 흐려지는 데서 생깁니다. CMDSPACE 는 세 영역을 하나의 설계로 묶습니다.

01 / Alignment

경영진과 현장을 정렬

임원 교육·코칭으로 조직의 실제 문제와 스폰서, 성공 기준을 먼저 확인합니다.

02 / Context

지식과 업무 흐름을 구조화

문서·회의·경험을 Context Pack과 검색 가능한 지식 구조로 전환해 AI가 일할 맥락을 만듭니다.

03 / Operations

책임과 품질을 운영 리듬으로

접근 권한·갱신 주기·관리자 역할·품질 기준을 문서화해 교육 뒤에도 계속 작동하게 합니다.

실제 화면 | 촬영 2026-09-10T22:07:51+09:00 | 공개 페이지의 Positioning 영역만 선택 캡처. 고객별 성과를 회사 성과로 재귀속하지 않음.

- 임원 교육과 코칭으로 조직의 실제 문제를 확인합니다.
- 경영진과 현장이 함께 볼 성공 기준을 정합니다.
- 도구 설명을 지식 구조와 운영 책임으로 연결합니다.

공개 서비스 페이지의 경영진 정렬, 지식 구조화, 운영 책임 세 영역입니다. 고객명이나 개별 코칭 내용은 포함하지 않습니다.

<https://cmdspace.work/#value>



지식운영 진단과 파일럿

한 부서의 실제 업무로 AI 지식운영의 조건을 확인합니다.

ENGAGEMENT MODEL

발견에서 운영까지, 다섯 단계로 연결합니다

고객의 문제·보안·데이터 성숙도에 따라 단계를 조정하며, 파일럿의 사용과 지표가 확인될 때만 확산을 권합니다.

01 Discover 문제 발견 임원 교육·코칭으로 우선 질문과 내부 스폰서를 확인합니다.	02 Diagnose 기준선 진단 지식 원천·핵심 업무·보안 조건·성공 기준을 정의합니다.	03 Pilot 부서 파일럿 한 부서·한 업무에 Context Pack·LLM Wiki·AI 흐름을 구축합니다.	04 Scale 조직 확산 수준별 핸드온·관리자 양성·전달체계로 반복 가능성을 만듭니다.	05 Operate 운영 정착 책임·갱신·품질·보안·분기 감사를 운영리듬으로 만듭니다.
--	--	--	---	--

실제 화면 | 촬영 2026-09-10T22:07:53+09:00 | 공개 서비스 모델의 상단만 크롭. 하단 인원 및 과거 강의 횟수 통계는 제외.

- 지식 원천, 핵심 업무, 보안 조건과 성공 기준을 진단합니다.
- 한 부서와 한 업무에 Context Pack(업무 맥락 묶음)과 LLM Wiki를 적용합니다.
- 파일럿 사용과 지표를 확인한 뒤 확산과 운영을 설계합니다.

문제 발견 → 기준선 진단 → 부서 파일럿 → 조직 확산 → 운영 정착의 공개 참여 모델입니다.

<https://cmdspace.work/#program>



Obsidian 에디터

생각을 읽고 고치는 작업 공간을 직접 운영합니다.

21

포맷은 사고를 강제한다

22

23

Marshall McLuhan의 유명한 문장인 “The medium is the message”**는 내용보다 매체의 형식 자체가 인간의 지각과 사회적 관계를 더 깊게 바꾼다는 주장이다. 사람들은 보통 메시지의 내용에 주목하지만, McLuhan은 오히려 **어떤 형식으로 전달되는가가 인간의 감각, 속도, 관계 방식, 사고 습관을 재구성한다고 봤다.

24

25

이 관점을 오늘의 언어로 더 실용적으로 바꾸면 이렇게 말할 수 있다.

26

27

포맷은 사고를 강제한다.

28

29

왜 포맷이 중요한가

30

같은 내용이라도 어떤 포맷에 담기느냐에 따라 사람은 전혀 다르게 생각하게 된다.

31

32

• PPT는 생각을 슬라이드 단위로 자르게 만든다.

33

• 표는 비교하게 만든다.

34

• 체크리스트는 실행 순서를 강제한다.

35

• Markdown은 텍스트를 구조화하고 연결하게 만든다.

36

• YAML은 의미를 필드 단위로 선언하게 만든다.

37

• Git은 수정 이력과 버전을 남기게 만든다.

38

39

즉 포맷은 표현의 껍데기가 아니다. 포맷은 무엇을 먼저 보게 만들고, 무엇을 생략하게 만들며, 무엇을 반복 가능하게 만드는가를 결정한다.

40

- 마크다운 본문에 주장과 근거를 함께 씁니다.
- 헤딩과 목록으로 사고의 구조를 드러냅니다.
- 완성 문서와 작업 중인 노트를 같은 지식 기반에서 연결합니다.

작성자의 실제 Obsidian에서 「포맷은 사고를 강제한다」를 연 에디터입니다. 줄 번호, 제목과 본문이 보입니다. 공개 URL은 교육 주제 소개이며 개인 볼트 접속 주소가 아닙니다.

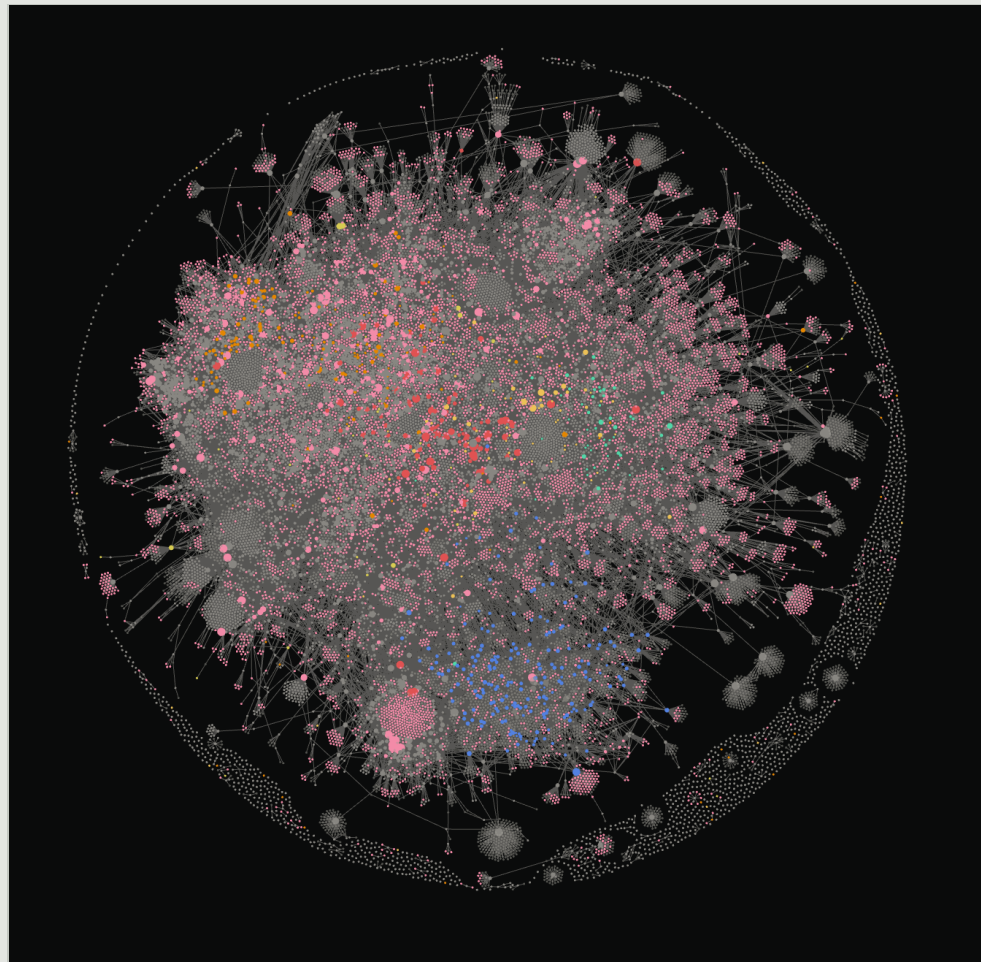
실제 화면 | 촬영 2026-09-10 | Obsidian 사용 예시 / 구요한의 실제 작업환경

<https://cmdspace.work/#topics>



Obsidian 그래프

노트 사이의 연결을 눈으로 확인합니다.



실제 화면 | 촬영 2026-09-10 | Obsidian 사용 예시 / 구요한의 실제 작업환경

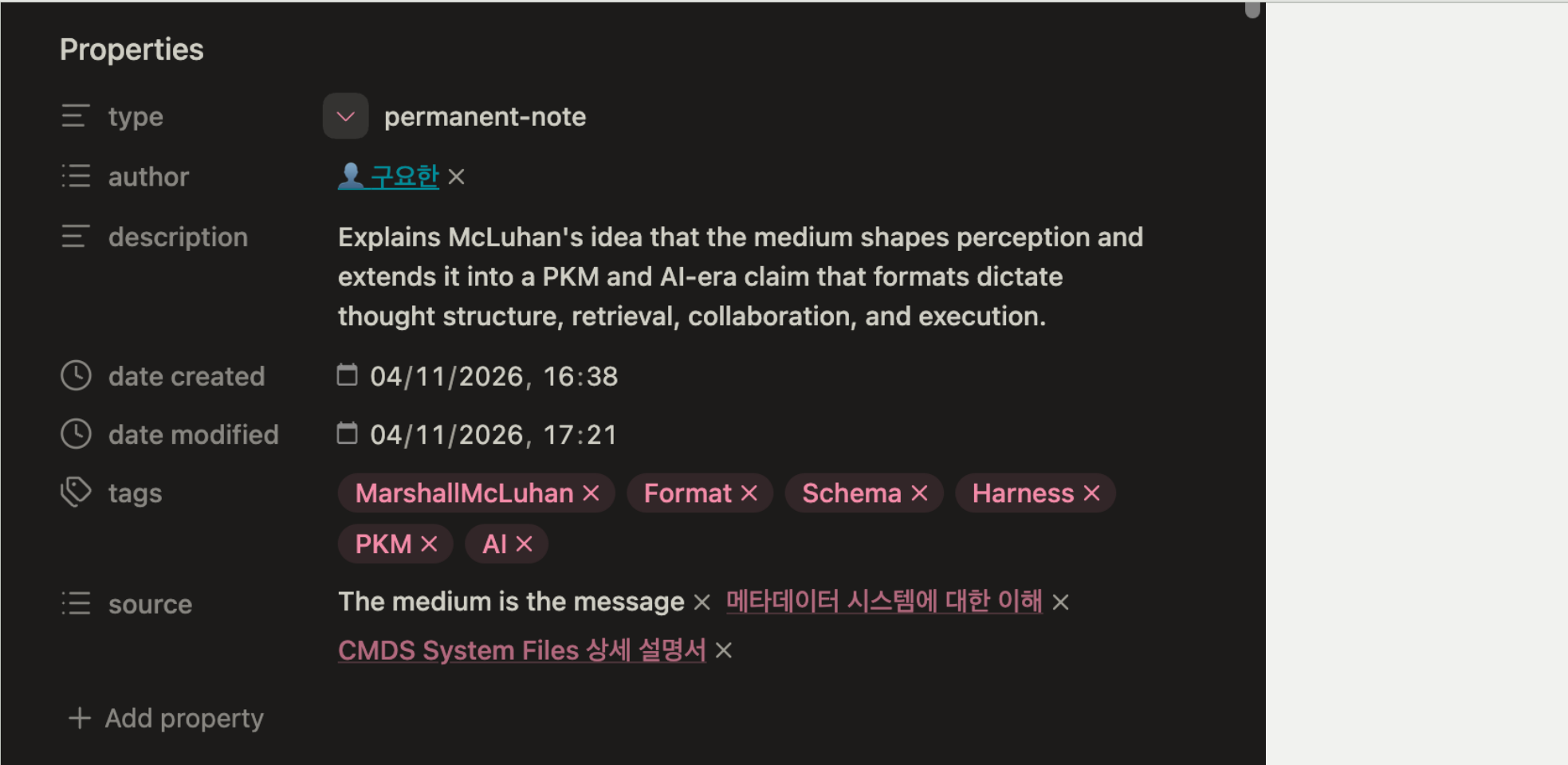
- 위키링크가 노트 사이의 실제 연결을 만듭니다.
- 전체 연결망을 살펴보고 필요하면 검색 필터로 범위를 좁힙니다.
- 그래프는 연결 탐색용 화면이며 지식의 질을 점수화한 결과가 아닙니다.

작성자의 실제 Obsidian 전체 그래프입니다. 노트명과 태그명을 숨기고 실제 노드와 연결선만 보여줍니다. 검색 필터 없이 렌더링한 지식망의 형태이며 노드 수나 지식의 질을 측정한 결과가 아닙니다. 개인 볼트 원문에 접근할 수 있는 화면은 아닙니다.

<https://cmdspace.work/#topics>

Obsidian 메타데이터

내용뿐 아니라 의미와 출처를 필드로 남깁니다.



실제 화면 | 촬영 2026-09-10 | Obsidian 사용 예시 / 구요한의 실제 작업환경

- type은 어떤 종류의 노트인지 알려줍니다.
- description과 tags는 검색과 재사용의 단서가 됩니다.
- 작성일과 출처는 근거를 다시 확인할 길을 남깁니다.

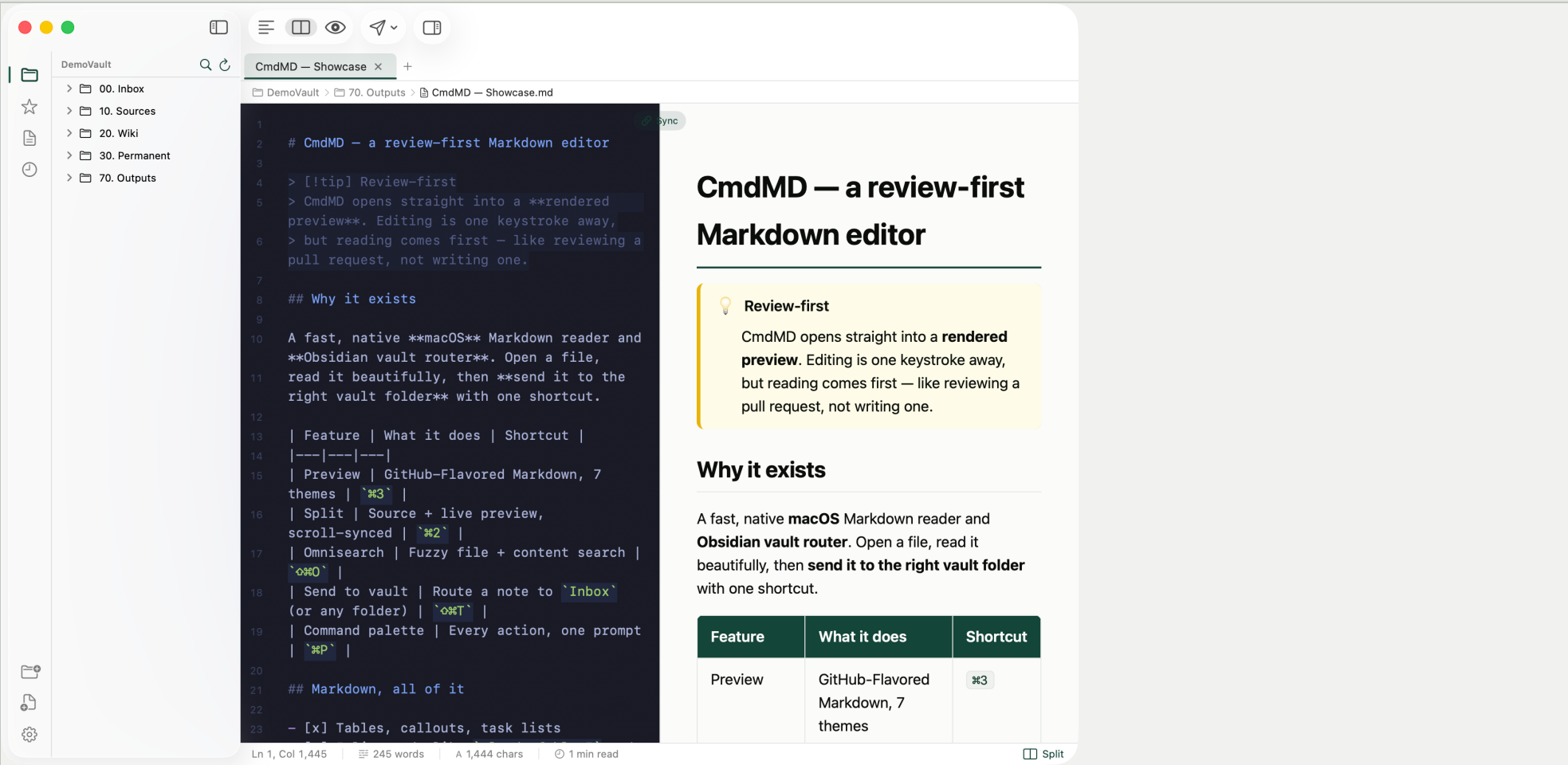
「포맷은 사고를 강제한다」의 실제 Properties 영역입니다. 노트 종류, 설명, 날짜, 태그와 출처가 표시됩니다. 이 화면은 모든 노트의 스키마 준수를 인증하지 않습니다.

<https://cmdspace.work/#topics>



CmdMD

마크다운을 먼저 읽고 필요한 곳으로 보냅니다.



실제 화면 | 촬영 미상 | 실제 앱 화면(촬영일 미상), 현재 UI와 다를 수 있음

- 렌더링 미리보기로 시작하는 macOS 마크다운 에디터입니다.
- 원문, 분할, 미리보기 화면을 전환합니다.
- 템플릿과 라우팅 규칙으로 Obsidian 볼트에 보냅니다.

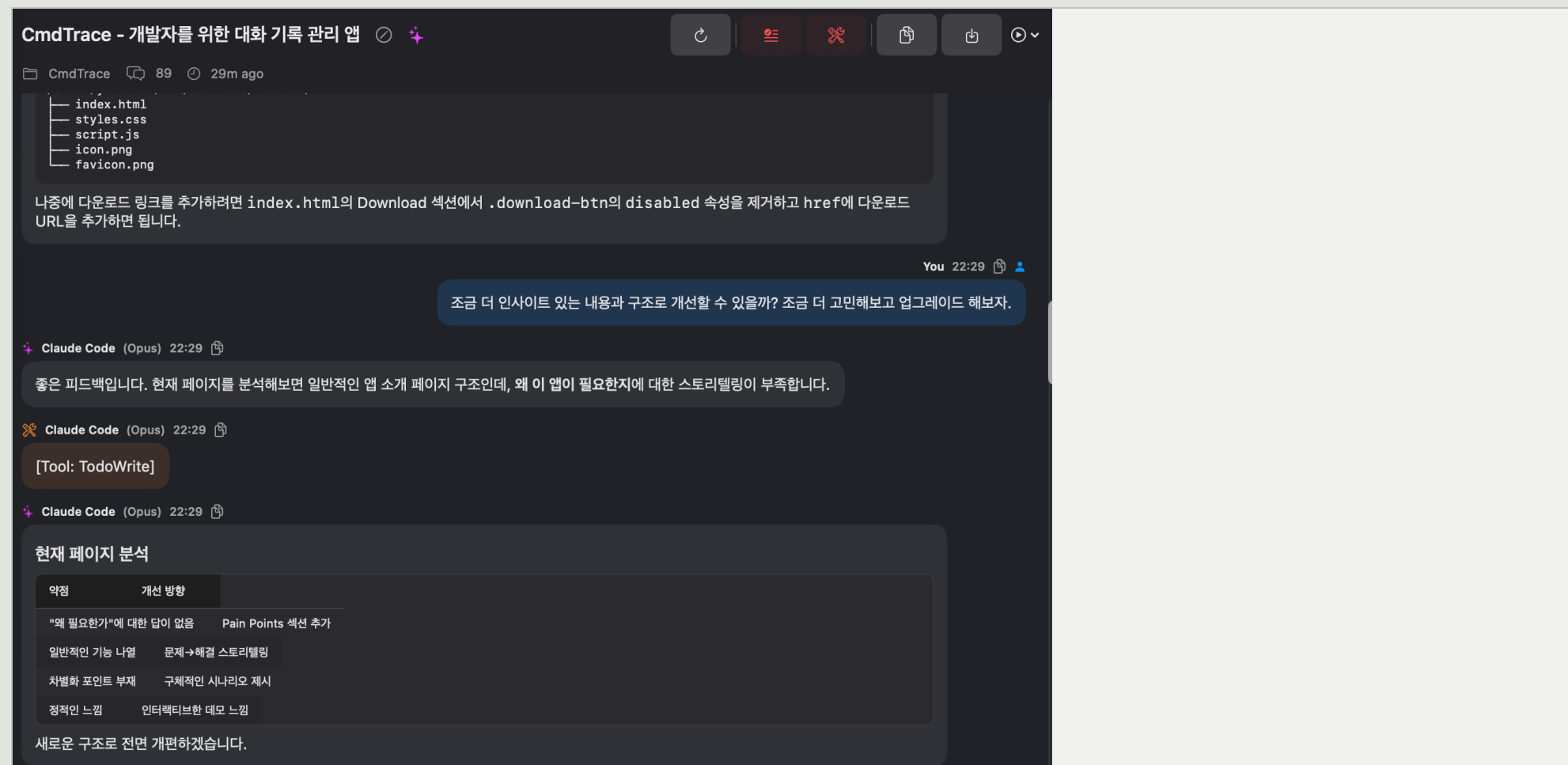
실제 앱의 DemoVault 예제입니다. 왼쪽은 파일 목록, 가운데는 원문, 오른쪽은 렌더링 미리보기입니다. 과거 공개 제품 캡처이며 현재 버전 번호를 뜻하지 않습니다.

<https://cmdmd.cmdspace.work>



CmdTrace

지난 AI 작업을 다시 찾고 이어갈 수 있게 합니다.



실제 화면 | 촬영 미상 | 실제 앱 화면(촬영일 미상), 현재 UI와 다를 수 있음

- AI 코딩 세션을 읽고 검색하는 macOS 앱입니다.
- 대화와 도구 실행 기록을 시간 흐름 안에서 봅니다.
- 흩어진 작업 기록을 다시 확인하는 인터페이스를 제공합니다.

CmdTrace 자체 소개 페이지를 개선하던 과거 세션을 실제 앱으로 읽는 화면입니다. 대화와 도구 실행 구분이 보입니다. 현재 세션이나 연구 원응답을 캡처한 것이 아닙니다.

<https://cmdtrace.cmdspace.work>

System Files

사람과 AI가 함께 따를 운영 규칙을 공개합니다.

^P
1

CLAUDE.md AI

Claude Code technical guide — file rules, vault commands, Obsidian syntax, output paths.

→ Claude Code

^P
2

AGENTS.md AI

Portable rules for every other AI coding agent operating in the vault.

→ Gemini · Codex · Cursor

^P
4

CMDS.md AI

System philosophy and user context — the why and the who.

→ All LLMs

^P
5

CMDS Guide HUMAN

Operational standards — required properties, templates, naming conventions.

→ User + AI

^P
6

CMDS Head Quarter HUMAN

The navigation hub — 91 categories across 100–900 CMDS sections.

→ User

^P
9

DESIGN.md VISUAL

Visual language spec — brand tokens, anti-AI-slop rules, skill-to-surface mapping.

→ All visual-output agents

실제 화면 | 촬영 2026-09-10T22:22:27+09:00 | 공개 배포 페이지 화면. 버전.분류 수는 정본 확인

- 기술 규칙, 시스템 철학, 표준과 탐색 지도를 문서로 나눕니다.
- 파일별 역할과 우선순위를 명시합니다.
- 자신의 지식관리 환경에 맞게 규칙을 가져와 조정할 수 있습니다.

공개 배포 대상 여섯 시스템 파일의 역할 카드입니다. 파일의 제목과 독자가 구분됩니다. 개인용 비공개 파일 본문은 포함하지 않습니다.

<https://system.cmdspace.work>



LLM Wiki

외부 자료를 다시 찾을 수 있는 연결된 지식으로 컴파일합니다.

3 운영 작업

Ingest · Query · Lint

위키의 호흡.

위키는 살아있는 시스템 — 들숨(Ingest)으로 raw source 를 흡수, 날숨(Query)으로 종합된 답변 생성, 위생점검(Lint)으로 일관성 유지. 마더십의 CMDS Process (Connect→Merge→Develop→Share) 와 다른 vocabulary 지만 보완적입니다.

 **/ingest**

Slash command · raw → wiki

외부 자료 (article · paper · book · transcript) 를 수집해 10. Raw Sources/ 에 저장 후 LLM 이 추출한 개념·엔티티 를 20. Wiki/ 에 컴파일. 수집 목적 (collectionPurpose) 1회 질문 후 진행.

 **/query**

Slash command · wiki → answer

qmd MCP (BM25 + Vector + HyDE) 로 위키와 raw sources 검색 후 답변 합성. wiki-worthy 면 30. Queries/ 에 file-back. 마더십 볼트도 함께 검색 가능 (cross-vault).

 **/lint**

Slash command · health check

Read-only 위생 점검 — orphan / contradiction / stale / frontmatter 일관성. Living Reference Policy 위반 (Core Context 30일 경과) 도 flag. 자동 수정 안 함, 사용자 결정 대기.

실제 화면 | 촬영 2026-09-10T22:07:57+09:00 | 공개 Operations 영역 선택 캡처. 파일 수나 현재 위키 내용은 공개하지 않음.

- 원문을 보존하고 별도의 위키에 지식을 정리합니다.
- Ingest(수집과 컴파일), Query(질의), Lint(위생 점검)를 구분합니다.
- 사람의 노트와 LLM이 정리한 참고 지식의 역할을 나눕니다.

공개 LLM Wiki 사이트의 세 운영 작업 설명입니다. 원문에서 위키로, 위키에서 답변으로 이어지는 역할을 인터넷 없이 설명할 수 있습니다.

<https://llm-wiki.cmdspace.work>



CMDS 스타터킷

빈 폴더가 아니라 운영 규칙이 있는 볼트로 시작합니다.

cmds-vault

Class-ready Obsidian vault running [CMDs conventions](#) with [Gobi Desktop](#) integration baked in. Clone, open in Obsidian, start writing.

Template version: 1.2.2 (released 2026-07-13). To check what version you're running: `cat VERSION` or read `template-version` in this README's frontmatter. Upstream changes are tracked in [CHANGELOG.md](#). Authoritative version source: git tags on this repo.

What this is

This vault is a **graft**:

- **Base** — [cmds-system-files](#) by John Koo. The 6 system files (`CLAUDE.md` , `AGENTS.md` , `CMDS.md` ,  `CMDS Guide.md` ,  `CMDS Head Quarter.md` , `DESIGN.md`) and 8 shared rules (`.claude/rules/`) plus 8 slash commands (`.claude/commands/`) define the CMDs PKM operating model — atomic notes, frontmatter standard, 4-stage pipeline (Connect → Merge → Develop → Share), wikilink discipline.
- **Add-on** — minimal Gobi identity files (`BRAIN.md` , `BRAIN.jpg` , `BRAIN_PROMPT.md`) so the vault registers as a Gobi Brain. Skills and prompts that drive Gobi live in cmds-style `90. Settings/91. Skills/` and `90. Settings/92. Prompts/`. The `.gobi/` runtime folder is **not** shipped — `gobi init` creates it on first run.

Numeric folders (`00. Inbox/` ... `90. Settings/`) are pre-created with `.gitkeep` placeholders matching `cmds-system-files/rules/directory-structure.md`. No manual setup required.

Quickstart

1. Clone and open

```
git clone https://github.com/johnfkoo951/cmds-vault.git ~/Documents/cmds-vault
```

실제 화면 | 촬영 2026-09-10T22:26:19+09:00 | 공개 README 영역 크롭. 저장소 별과 다운로드 수 제외. README 내부 날짜 표기가 서로 달라 최신 버전이나 배포일은 본문에서 단정하지 않음.

- CMDS 폴더 구조와 기본 규칙을 포함한 Obsidian 템플릿입니다.
- Connect → Merge → Develop → Share 흐름을 따라 시작합니다.
- 자신의 용도에 맞게 내용을 채우고 규칙을 조정합니다.

공개 GitHub 저장소의 README 소개와 구성 설명입니다. 사용자 개인 볼트가 아닌 배포용 스타터킷입니다.

<https://github.com/johnfkoo951/cmds-vault>



AKM Index

노트 수가 아니라 작동하는 지식관리 시스템을 점검합니다.

P	Prompt Engineering20% 에이전트에게 주는 지시가 버전 관리되는 자산인가 — 계층화 · 계약 명세 · 트리거 라우팅 · 예시 · 수명주기. → P1~P5
C	Context Engineering25% 에이전트가 맞는 지식을 적정 분량으로 찾을 수 있는가 — 주소체계 · 메타데이터 · 검색 · 점진 공개 · 출처와 신선도. → C1~C5
H	Harness Engineering20% 런타임이 강력하고 안전하게 정비됐는가 — 도구 표면 · 스킴 라이브러리 · 가드레일 · 지속 메모리 · 오케스트레이션. → H1~H5
L	Loop Engineering20% 사이클이 실제로 돌고 복리로 쌓이는가 — 파이프라인 · 주기 리뷰 · 적체 관리 · 복리 환류 · 자기교정. 최근 30일 가동 증거 필수. → L1~L5
X	Interop & Governance15% 하나의 런타임·디스크 너머에서도 살아남는가 — 런타임 이식성 · 토폴로지 · 복구 · 보안 · 전수 가능성. → X1~X5

실제 화면 | 촬영 2026-09-10T22:22:42+09:00 | 공개 필러 카드만 재캡처. 참여자 평가 보드와 개인별 점수 제외.

- 프롬프트, 컨텍스트, 하네스, 루프, 상호운용성과 거버넌스를 살펴봅니다.
- 다섯 필러와 25개 기준에 행동과 증거를 대응합니다.
- 자기보고와 실제 파일, 설정, 날짜 있는 기록을 구분합니다.

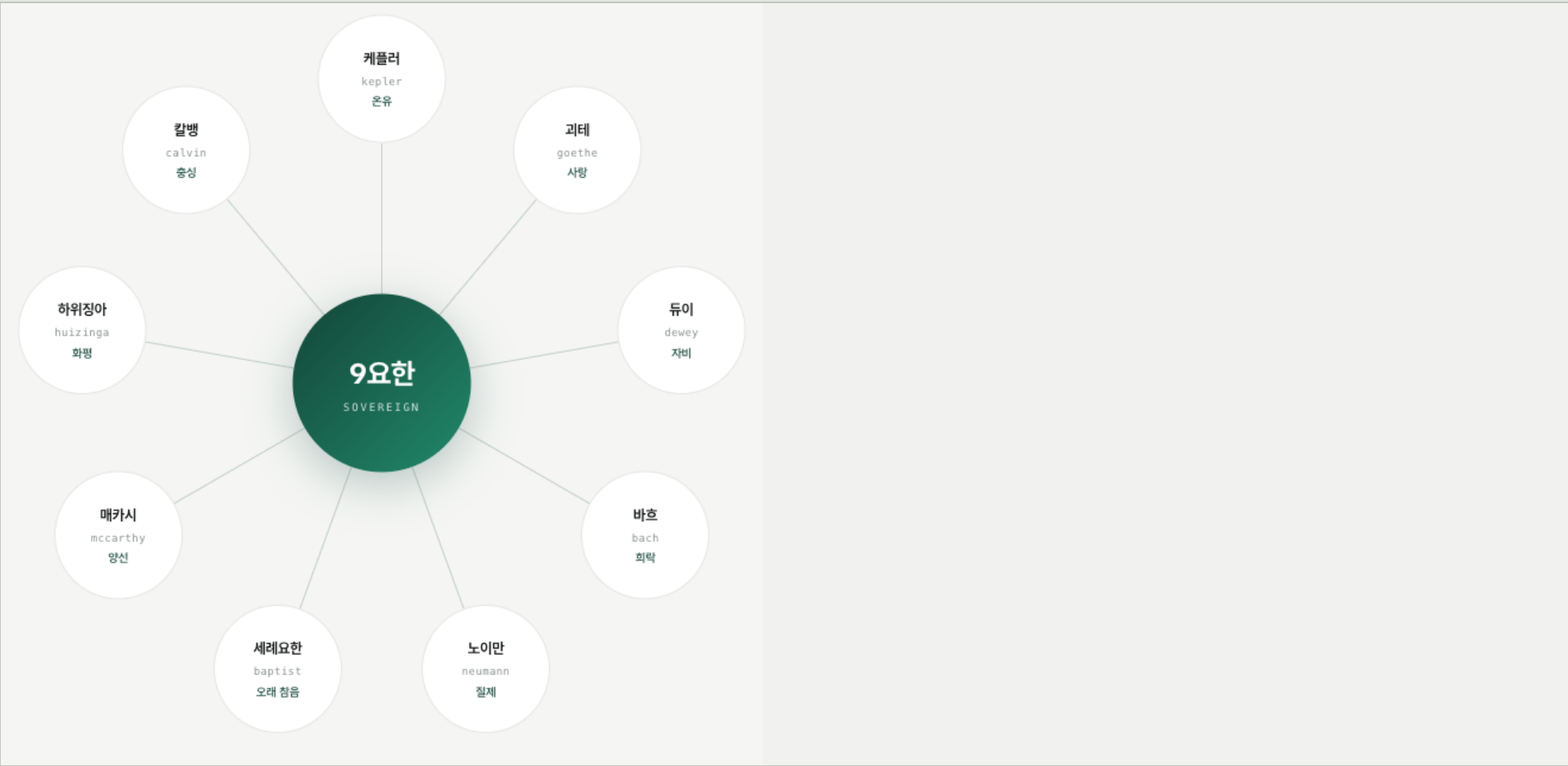
공개 루브릭의 다섯 평가 필러입니다. 각 필러가 지시, 지식 공급, 실행 환경, 반복 개선, 이식성과 운영 책임을 구분합니다. 임의 성능 인증이나 독립 검증 완료를 뜻하지 않습니다.

<https://akm.cmdspace.work>



9Yohan

전문 역할을 나누되 최종 판단과 조정은 하나로 모읍니다.



실제 화면 | 촬영 2026-09-10T22:22:45+09:00 | 공개 Star Diagram만 선택 캡처. 사이트 자체의 구조 그림이며 새로 생성한 UI가 아님.

- 연구, 글쓰기, 교육, 창작 등 아홉 역할을 구분한 에이전트 설계입니다.
- 중앙 조정자가 요청을 나누고 결과를 합칩니다.
- 전문가 간 역할 분리와 최종 승인 책임을 명시합니다.

공개 아키텍처 사이트의 중앙 조정자와 아홉 전문 역할 연결도입니다. 실제 사이트 화면의 설계 설명이며 실시간 에이전트 가동 상태를 보여주는 대시보드가 아닙니다.

<https://9yohan.cmdspace.work>

CMDSPACE Apps

직접 만든 도구를 한곳에서 탐색합니다.



실제 화면 | 촬영 2026-09-10T22:08:26+09:00 | 공개 Hero의 제목과 설명만 크롭. 앱 수, 라이브 수와 기타 변동 통계 제외.

- macOS 앱, Obsidian 플러그인, 웹 플랫폼과 개발 도구를 모읍니다.
- 제품별 소개와 공개 링크를 따라갈 수 있습니다.
- 목록 등재와 실제 출시 상태는 구분해서 확인합니다.

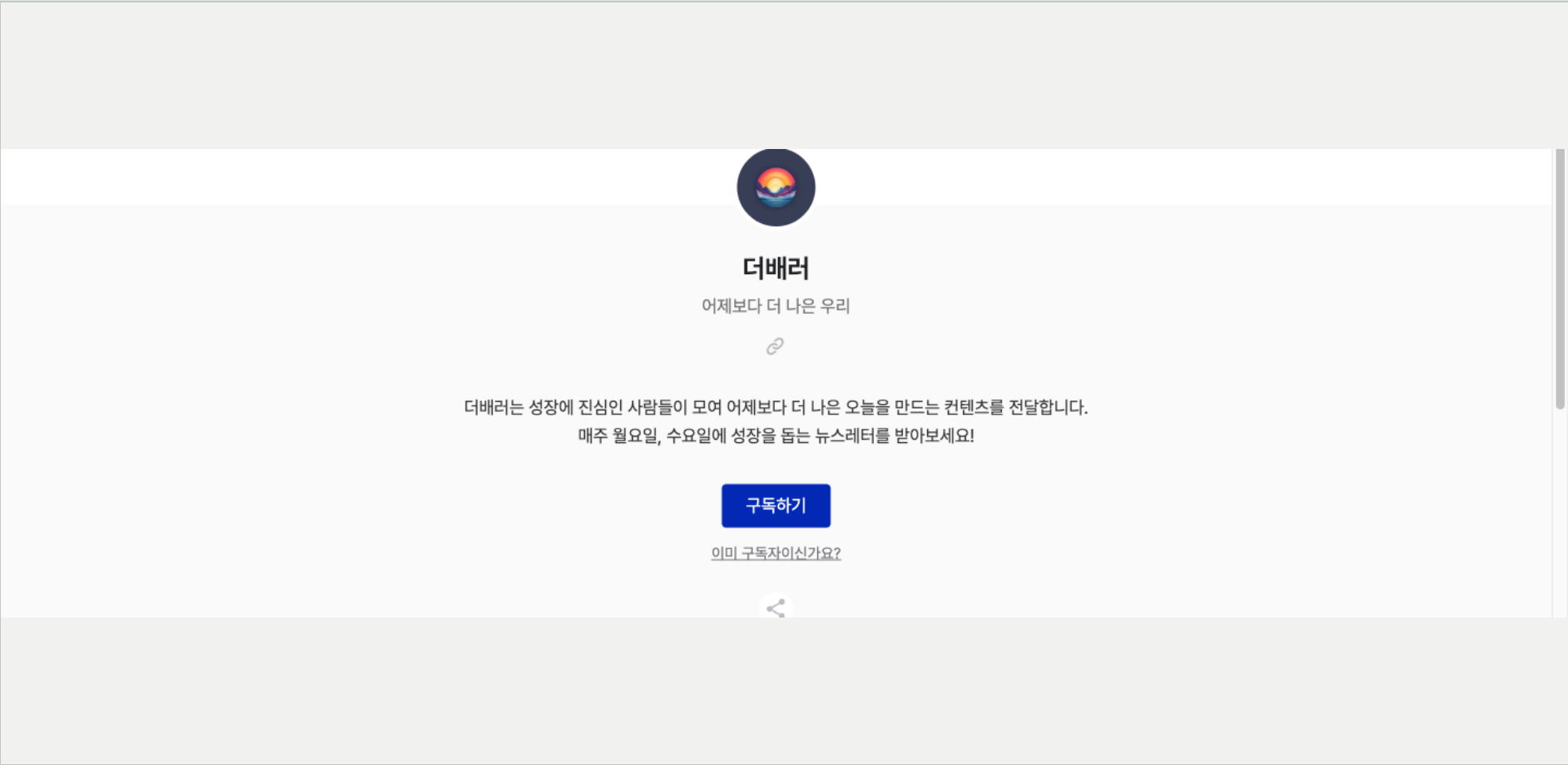
공개 Apps 카탈로그의 소개 화면입니다. 전체 앱 수와 라이브 수는 빠르게 달라지므로 크롭에서 제외했습니다.

<https://apps.cmdspace.work>



더배러

성장과 지식 활용의 생각을 뉴스레터로 나눕니다.



실제 화면 | 촬영 2026-09-10T22:26:23+09:00 | 뉴스레터 아카이브 화면. 발송 일정은 채널에서 확인

- 어제보다 더 나은 우리를 주제로 콘텐츠를 전합니다.
- 공개 아카이브에서 지난 글을 읽을 수 있습니다.
- 구독 페이지에서 이후 발행 소식을 받을 수 있습니다.

더배러의 실제 공개 뉴스레터 소개와 구독 화면입니다. 최근 글 목록은 제3자 이메일이 표시된 카드가 있어 포함하지 않았습니다.

<https://thebetter.stibee.com/>

월간옵시디언

실제 지식관리 경험을 지속적으로 함께 배우는 자리입니다.

Classes Intro

Curriculum

Class Reviews

Notices

003

UNLOCK THE POWER OF YOUR MIND

CMDSPACE

월간 옵시디언

세미나는 매월 1회, 온라인과 오프라인에서 진행합니다

[상시 모집] 월간 옵시디언(정기 멤버십)

실제 화면 | 촬영 2026-09-10T22:15:26+09:00 | 커뮤니티 소개 화면. 모집·다음 일정은 링크에서 확인

- Obsidian과 지식관리 주제를 정기적으로 다룹니다.
- 도구 활용과 AI 연결 사례를 함께 살펴봅니다.
- 참여 조건과 최신 일정은 공식 안내에서 확인합니다.

공식 클래스 사이트의 월간옵시디언 멤버십 소개 이미지와 제목입니다.
결제 정보, 개인 수강 내역, 지난 일정을 다음 일정처럼 보여주는
영역은 제외했습니다.

<https://pbl.to/MrSU8P>

CMDSPACE Share

노트를 웹으로 공유하고 공유 이후도 관리합니다.

CMDS Share

English · 한국어

Share Obsidian notes to the web with one command — and stay in control after you hit share.

Unlike simple publish plugins, CMDS Share is **governance-first**: shared notes carry server-side view counts, expiry, and one-click revocation, and a web dashboard shows everything you have published. Notes are rendered through Obsidian's own pipeline, so tables, callouts, embedded images, and your theme look exactly like they do in your vault.

Part of the [CMDSPACE](#) ecosystem.

Features

- **One-command publishing** — `Share current note to web` copies a public URL to your clipboard
- **Faithful rendering** — Obsidian's real renderer (tables, callouts, code, Korean/CJK) + your theme CSS
- **Share page extras** — table of contents sidebar with scrollspy, an interactive local graph (force-directed, zoom/pan/hover-highlight), code-block copy buttons, heading deep links, dark/light mode. TOC and graph can float as popups or dock as a persistent sidebar
- **Governance** — server-side registry, view counts (link-preview bots filtered), expiry with a tombstone page, revoke/restore without deleting, cross-vault ID protection
- **End-to-end encryption (optional)** — AES-256-GCM in the plugin; the key travels only in the URL fragment and never reaches the server
- **CMS view in Obsidian** — browse all shares, live view counts reconciled from the server, re-share, revoke, delete
- **Web dashboard** — manage your shares from any browser
- **Five backends** — pick the hosting model that fits you (see below)

실제 화면 | 촬영 2026-09-10T22:26:21+09:00 | 공개 README 기능 영역만 크롭. 개인 공유 목록, 인증정보 및 비공개 서버 정보 없음.

- Obsidian 노트를 공개 링크로 공유하는 플러그인입니다.
- 공유 페이지에 목차와 로컬 그래프를 제공합니다.
- 만료, 철회와 선택적 암호화 지원은 사용 백엔드의 기능을 확인해야 합니다.

공개 저장소의 기능 설명입니다. 실제 관리 대시보드나 고객 공유 목록이 아니라 공개 README를 캡처했습니다. CMDSPACE 운영 서버를 모든 사용자가 쓸 수 있다는 뜻은 아닙니다.

<https://github.com/johnfkoo951/cmds-share>



자료를 다시 보고 다음 질문을 이어가세요

구요한 | CMDSPACE | johnfkoo951@gmail.com

현재 발표자료



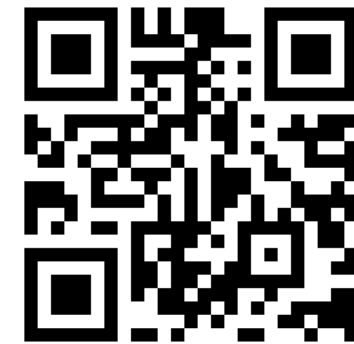
<https://labs.cmdspace.work/stepi-research-overview/>

CMDSPACE



<https://cmdspace.work>

구요한 프로필



<https://bio.cmdspace.work>