DGC

GroundProof 소개서

GroundProof 데이터 솔루션

보정과 검증을 위한 소비자 실험 데이터

고객사가 설계한 실험을 실제 시장에서 집행하고,
배정 기록과 원자료를 돌려드립니다.



Digicon Holdings Pte. Ltd. · Singapore · 2026년 11월

GroundProof 한눈에 보는 실험 채널

누가 쓰고, 무엇을 하고, 무엇을 받는지, 그리고 어떤 원칙을 지키는지입니다.

대상

소비자 반응을 예측하는 AI 기업과 연구기관

하는 일

고객사가 설계한 실험을 실제 시장에서 집행하고 측정

받는 것

원자료 · 배정 기록 · 노출당 지표

원칙

평가는 고객사가 직접, 결과는 비공개

세 가지 상품

맞춤 실험

고객사가 지정한 요인 조합을 독립 소재로 제작해 무작위로 게시

홀드아웃 데이터

모델이 본 적 없는 새 메시지를 대조군과 함께 실측한 원자료

구매 연결 실험

행동 기록과 주문 결과를 한 쌍으로 결합

디지콘은 **실험을 집행**하고, 고객사는 그 데이터로 **모델을 보정하고 검증**합니다.

사람 행동을 시뮬레이션하는 AI의 부상

사람의 응답과 행동을 모사하는 모델에 자본이 몰리고 있습니다.

Simile

20억 달러

기업 가치 · 2026.7

Series A · B로 2026년에만 3억 달러 이상

Ground Truth Gallup 패널 인터뷰

Aaru

10억 달러

헤드라인 기업 가치 · 2025.12

Series A · 누적 투자 약 8,800만 달러

Ground Truth EY 투자자 설문

Electric Twin

1,000만 달러

최근 투자 · 2026.2

누적 1,400만 달러 · 기업 가치 비공개

Ground Truth News UK 독자 패널

세 기업의 Ground Truth	설문 · 패널 · 인터뷰
실제 광고를 거친 행동 기록	공개 거래 사례 없음

행동을 예측하는 AI에 자본이 몰리지만, **실제 시장 행동의 Ground Truth**는 아직 드뭅니다.

출처: Simile(2026.7.30), TechCrunch(2025.12.5), PitchBook(Fortune 2026.6.17 인용), Electric Twin(2026.2.11), Gallup(2026.5.11), EY(2026.9.30), Research Live(2026.5.1), Neudata(2025.6.10). Ground Truth: 모델의 예측을 평가하는 실측 정답 데이터.

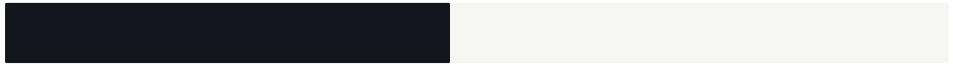
실제 행동과의 대조를 요구하는 연구 결과

언어 모델은 설문 실험에 강하고, 실제 행동을 재는 현장 실험에 약합니다.

Hewitt 외 · Stanford · 2024 프리프린트 기준

한 연구 안의 여러 메시지 효과를 언어 모델이 얼마나 맞혔는가 · 상관계수

설문 실험 메가스터디



0.47

현장 실험 메가스터디 · 실제 행동



0.27

같은 시험의 문제별 점수를 맞히는 일과 같습니다. 설문지 시험보다 실전 시험의 문제별 점수를 맞히기가 훨씬 어렵고, 1에 가까울수록 정확합니다.

참고 · 설문 실험 70건의 실험 간 효과 비교에서는 0.85

Park 외 · 2024.11

1,052명과의 2시간 인터뷰로 만든 에이전트의 응답 재현 수준이, 본인이 2주 뒤 다시 답한 일관성의 85%에 이르렀습니다.

Kutzner 외 · 2026.9

사람 행동 예측 모델의 검증 기준은 진술보다 관측된 행동이어야 한다고 제시했습니다.

NeurIPS 2025 · ICML 2026

「소셜 월드 모델」이라는 표현이 학회 워크숍과 본 학회 논문에 등장했습니다.

AI는 **설문의 답**보다 **시장의 행동**을 맞히기 어렵고, 그 간극을 메우는 것이 **실측 기록**입니다.

출처: Hewitt · Ashokkumar · Ghezae · Willer(2024 프리프린트, 이후 Nature 게재), Park 외(2024.11), Kutzner 외(2026.9), Zhou 외(NeurIPS 2025 워크숍), Yu 외(ICML 2026). 메가스터디 수치는 연구별 상관의 메타 분석 평균(보정 전)입니다.

GroundProof 실제 행동이 빠진 예측 모델

설문을 잘 맞히는 AI도 실제 행동 앞에서는 정확도가 떨어집니다.

AI 예측과 실제 결과의 일치도

1에 가까울수록 정확합니다

설문 실험 모의고사에 해당



0.47

현장 실험 · 실제 행동 실전 시험에 해당



0.27

모의고사 점수로 실전 성적을 맞히기가 어려운 것과 같습니다.

무작위 배정이 필요한 이유

두 집단의 특성을 비교했을 때 차이가 큰 항목의 비율

Meta A/B 테스트 181,890건 · 광고별로 관객이 다름



22%

Meta Lift 테스트 3,204건 · 무작위 대조군



0.16%

공정하게 나뉘었다면 0에 가까워야 합니다. 심사위원마다 다른 손님을 붙이면 맛의 차이와 손님의 차이를 가릴 수 없는 것과 같습니다.

실제 시장에서 **무작위로 배정한 실험 데이터**가 그 간극을 메웁니다.

출처: Hewitt 외(2024 프리프린트, 메가스터디 기준 상관계수), Burtch 외(arXiv 2508.21251, 2025.8, 동료 심사 전). 오른쪽 막대는 25%를 끝으로 그렸습니다.

GroundProof 고객사가 설계하고 디지콘이 집행하는 실험

평가는 고객사가 직접 하고, 디지콘은 원자료와 배정 기록을 돌려드립니다.

실험 채널의 네 단계

01 실험 설계 고객사

자극과 요인을 지정하고, 예측은 고객사 안에서 먼저 확정

02 제작 · 게시 디지콘

독립 소재를 제작하고 무작위 배정으로 게시

03 측정 디지콘

테스트 랜딩 페이지에서 배정 기록과 행동을 측정

04 원자료 전달 고객사

배정 기록과 함께 전달하고, 평가는 고객사가 직접

상품 3종

맞춤 실험

고객사가 지정한 요인 조합을 독립 소재로 제작해 무작위로 게시합니다. 같은 대본을 여러 크리에이터가 수행하는 교차 실험을 포함합니다.

홀드아웃 데이터

모델이 본 적 없는 새 메시지를 대조군과 함께 실측한 원자료입니다.

구매 연결 실험

행동 기록과 주문 결과를 한 쌍으로 결합합니다. 주문 연결 동의가 갖춰진 과제에서 진행합니다.

결과는 **고객사만 보유**하고, 디지콘은 **공개하지 않습니다**.

GroundProof 실험 채널의 세 가지 상품

고객사가 실험을 설계하고, 디지콘이 실제 시장에서 집행해 원자료를 돌려드립니다.

맞춤 실험

고객사가 지정한 요인 조합을 독립 소재로 제작하고, 계정과 게시 순서를 무작위로 배정합니다.

고객사의 몫
요인 지정 · 예측의 사전 확정

디지콘의 몫
제작 · 무작위 게시 · 배정 기록

받는 것
원자료 · 배정 기록 · 노출당 지표

같은 대본을 여러 크리에이터가 수행하는 교차 실험 포함

홀드아웃 데이터

모델이 본 적 없는 새 메시지를 테스트 랜딩 페이지에서 대조군과 함께 실측합니다.

고객사의 몫
예측의 사전 확정 · 직접 평가

디지콘의 몫
메시지 제작 · 무작위 배정 · 측정

받는 것
메시지별 행동 원자료 · 배정 확률

모든 실험군이 같은 광고로 유입

구매 연결 실험

행동 기록과 주문 결과를 한 쌍으로 결합합니다.

고객사의 몫
주문 연결 동의 · 판매 합계 제공

디지콘의 몫
구매 코드 발급 · 행동과 주문의 결합

받는 것
실험군별 행동과 주문의 쌍

주문 연결 동의가 갖춰진 과제에서 진행

평가는 **고객사가 직접** 하고, 결과는 **고객사만 보유**합니다.

고객사가 설계한 실험으로 만든 데이터의 재판매 제한과 이용 범위는 계약의 데이터 조항으로 정합니다.

GroundTruth 기록 실험 데이터 한 건의 구성

메시지, 유입과 배정, 비식별 속성, 실제 반응, 구매 신호를 한 건에 담습니다.

예시 · 실제 기록 아님

항목	예시 값
시장 · 카테고리	미국 · 스킨케어 클렌징밤
유입	정적 광고 1종 클릭 · 유입 경로 표시(유료 또는 오가닉)
배정	후보 메시지 C · 배정 확률 25% · 대조군 20%
카피	"Takes off sunscreen in one step." 기대
응답자 속성	25세에서 34세 · 현재 폼 클렌저 사용 · 비식별
반응	체류 48초 · 스토어 이동
응답 원문	"Worried it will leave my skin greasy." 마찰
구매 신호	후보 메시지별 자사몰 판매 합계(브랜드 동의, 개인 연결 없음)
판정 · 근거 수준	채택 · 대조군 설계

기록에 넣는 것

- 테스트 랜딩 페이지의 동의 응답
- 무작위 배정과 배정 확률
- 유입 경로(유료 · 오가닉)
- 브랜드가 서면 동의한 판매 합계
- 다섯 가지 구매 동인 태그
- 동의 범위 : 연구 · 모델 보정과 검증

기록에 넣지 않는 것

- 개인 식별 정보
- 구매자 개인 단위 주문 정보
- 플랫폼 원자료 그대로의 재판매

판정용 기록은 **테스트 랜딩 페이지**에서만 만들고, 오가닉 반응은 **집계 · 파생 등급**으로만 씁니다.

예시 값은 설명을 위한 가상 기록입니다. 기록 형식과 기록 설명서는 첫 셀 운영 전에 확정합니다. 셀: 한 시장의 한 카테고리를 묶은 운영 단위.

기존 데이터 원천과 실측 기록의 비교

기존 원천은 실제 반응, 무작위 배정, 라이선스 가운데 하나 이상이 빠집니다.

질문

실제 시장 반응, 메시지 단위 비교, 제3자 라이선스를 함께 갖춘 데이터 원천은 무엇인가

데이터 원천	실제 시장 반응	메시지 단위 무작위 배정	제3자 라이선스
설문 패널	× 진술 응답	○	○
플랫폼 A/B 테스트	○	△ 광고별 오디언스 상이	× 약관상 금지
구매 패널	○ 실구매	△ 캠페인 단위 연결	△ 사업자 보유
합성 응답	× 모델 추정	해당 없음	○
디지콘 (설계 기준)	○ 광고 클릭 방문자 범위	○	○ 동의 · 계약 충족 시

설문은 **「사시겠습니까」**라고 묻고, 테스트 랜딩 페이지는 실제 광고를 보고 들어온 사람이 **무엇을 눌렀는지** 기록합니다.

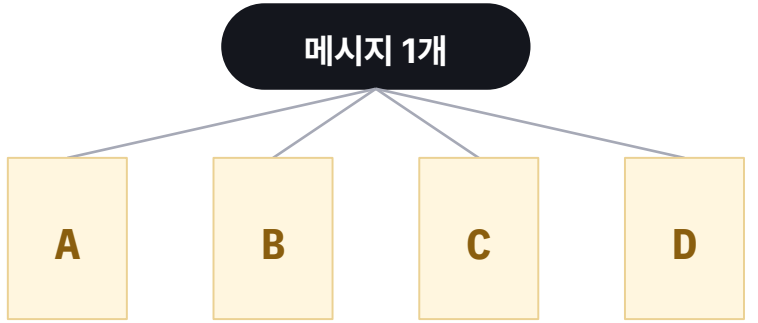
○ 충족 · △ 일부 충족 · × 충족하지 않음. 디지콘 행은 설계 기준이며, 동의 문구와 데이터 조항은 법률 검토 후 확정합니다. 근거: Meta Platform Terms(2026.2.3 개정), Burtch 외(2025.8), Circana NCSolutions, Kantar(2024.2.15)

영상이 높이는 반응의 해상도

영상은 노출 한 번마다 정지 이미지에 없는 시간 지표를 남깁니다.

신호	정지 이미지	영상
좋아요 · 공유 · 저장	기록	기록
평균 시청 시간	없음	기록 · Instagram 릴스
2초 · 6초 동영상 조회수	없음	기록 · TikTok 광고
25 · 50 · 75 · 100% 동영상 조회수	없음	기록 · TikTok 광고
첫 3초 건너뛰기 비율	없음	기록 · Instagram 릴스
재시청을 포함한 재생 시간	없음	기록 · TikTok 광고
참여 기반 조회수	없음	기록 · YouTube

같은 메시지, 여러 편의 독립 소재



같은 영상의 조회는 서로 독립된 관측으로 보기 어려우므로, 메시지의 힘은 여러 소재에 걸쳐 읽어야 영상 한 편의 우연과 구분됩니다. 한 사람에게 같은 질문을 백 번 묻는 것은 백 명에게 묻는 것과 다릅니다.

같은 채널 안에서 **한 메시지를 여러 편의 독립 소재로** 테스트하고, 게재빈도 상한을 둡니다.

출처: Instagram Platform API · 릴스 인사이트(2025.8), TikTok Ads Manager 도움말(2025.11), YouTube 고객센터, Judd · Westfall · Kenny(2017), Campbell · Keller(2003). 같은 광고를 반복 노출하면 효과가 약화됩니다. YouTube 조회수는 2026.8.24부터 첫 프레임 기준입니다.

다섯 구매 동인 소비자가 멈추는 다섯 지점

가장 가치 있는 페르소나가 멈춘 지점부터 진단합니다.

01

계기

카테고리를 삶에 들여온 순간

이론 근거

JTBD 네 가지 힘
밀기 · 끌기 · 불안 · 습관

「매일 선크림을 바르기 시작했어요」

02

신념

문제의 원인과 해결에 대한 믿음

이론 근거

수단-목적 사슬
속성 → 결과 → 가치

「기름은 기름으로 지워야 한다」

03

비교

실제로 저울질하는 대안과 기준

이론 근거

동등점 · 차별점
준거 가격

「폼 클렌저로 두 번 세안하면 되는데?」

04

기대

원하는 결과와 그 판단 방식

이론 근거

기대 불일치
기대 · 결과 · 정서

「선크림까지 한 번에 지워졌으면」

05

마찰

사고 싶은데 사지 못하는 제약

이론 근거

행동경제학 마찰
번거로움 · 지불의 고통 · 불확실성

「사고 싶은데 배송이 2주라서요」

공통 규칙 모든 발견을 통찰과 시장의 VOC 원문으로 함께 기록하고, 구매 후 평가가 다른 소비자의 계기가 되는 재구매 2회전까지 추적하도록 설계했습니다.

같은 제품도 **시장마다 멈춘 지점**이 다르고, 지점마다 **다른 메시지**가 필요합니다.

예시 문구는 미국 클렌징밤 시장을 가정한 설명용입니다. 이론 근거: Christensen 외(2016, JTBD), Gutman(1982), Keller 외(2002), Oliver(1980), Prelec · Loewenstein(1998).

데이터 원칙 지키는 것과 한계

데이터의 품질을 지키는 장치와, 데이터가 대표하지 않는 범위를 함께 밝힙니다.

지키는 것

무작위 배정
모든 실험군이 같은 광고로 들어오고, 테스트 랜딩 페이지에서 배정 확률을 기록합니다.

기준의 사전 확정
측정할 행동과 최소 표본을 집행 전에 문서로 정하고, 미달한 메시지는 보류로 기록합니다.

분류의 이중 검수
구매 동인 태그는 표본의 10%를 두 사람이 따로 분류해 일치도를 보고합니다.

결과 비공개
원자료는 고객사가 보유하고, 디지콘은 결과를 공개하지 않습니다.

한계

대표성
기록은 광고를 보고 들어온 방문자의 반응이며 확률 표본이 아닙니다. 추론 범위, 유입 경로, 제외 비율을 기록 설명서에 적습니다.

플랫폼 데이터
플랫폼 원자료는 재판매하지 않고, 오가닉 반응은 집계 · 파생 등급으로만 씁니다.

동의와 계약
동의 문구와 데이터 조항은 시장별 법률 검토 후 확정합니다.

데이터의 **신뢰**는 집행 전에 정한 **절차**에서 나옵니다.

분류 검수: 표본 10% 이중 분류, 유형별 일치도 보고. 근거: 사람의 검수 없이 모델 분류만 쓰면 가설 검정의 약 31%가 틀린 결론에 이르렀습니다(Baumann 외, arXiv 2509.08825, 2025, 상위 모델 기준).

운영 기반 대량 제작 체계와 세 거점

실험에 필요한 독립 소재를 대량으로 만들고, 현지어로 집행합니다.

월 2,000편

AI 숏폼 제작 체계

3개 거점

한국 영업 · R&D, 싱가포르 본사, 베트남 제작

AI의 실행, 사람의 판정

반복 작업은 AI 에이전트, 판정 기준과 검수는 사람

실증 일정

2026 Q4 — **첫 셀의 첫 검증 기록**
기록 형식과 기록 설명서 확정

2027 Q1 — **분류 신뢰도와 두 번째 셀**
구매 동인 분류의 평가자 간 일치도 0.80 이상

2027 H1 — **판정 규칙 공개**
신규 시장 · 카테고리 확장

같은 대본을 **여러 크리에이터와 계정**이 수행하는 실험은 **대량 제작 체계**가 있어야 가능합니다.

제작 규모는 자체 집계(2026.10 기준)입니다. 일정은 2026년 10월 기준 실증 계획입니다. 셀: 한 시장의 한 카테고리를 묶은 운영 단위.

시작 방식 맞춤 실험 1회차

실험할 자극과 집행 시장을 정해 시작합니다.

첫 프로젝트

고객사가 설계한 맞춤 실험 1회차

착수 전 확정 항목

- 실험할 자극과 요인
- 집행 시장과 기간
- 측정할 행동과 최소 표본
- 원자료의 이용 범위

이어지는 상품

홀드아웃 데이터 · 구매 연결 실험

첫 실험의 결과를 본 뒤 함께 정합니다.

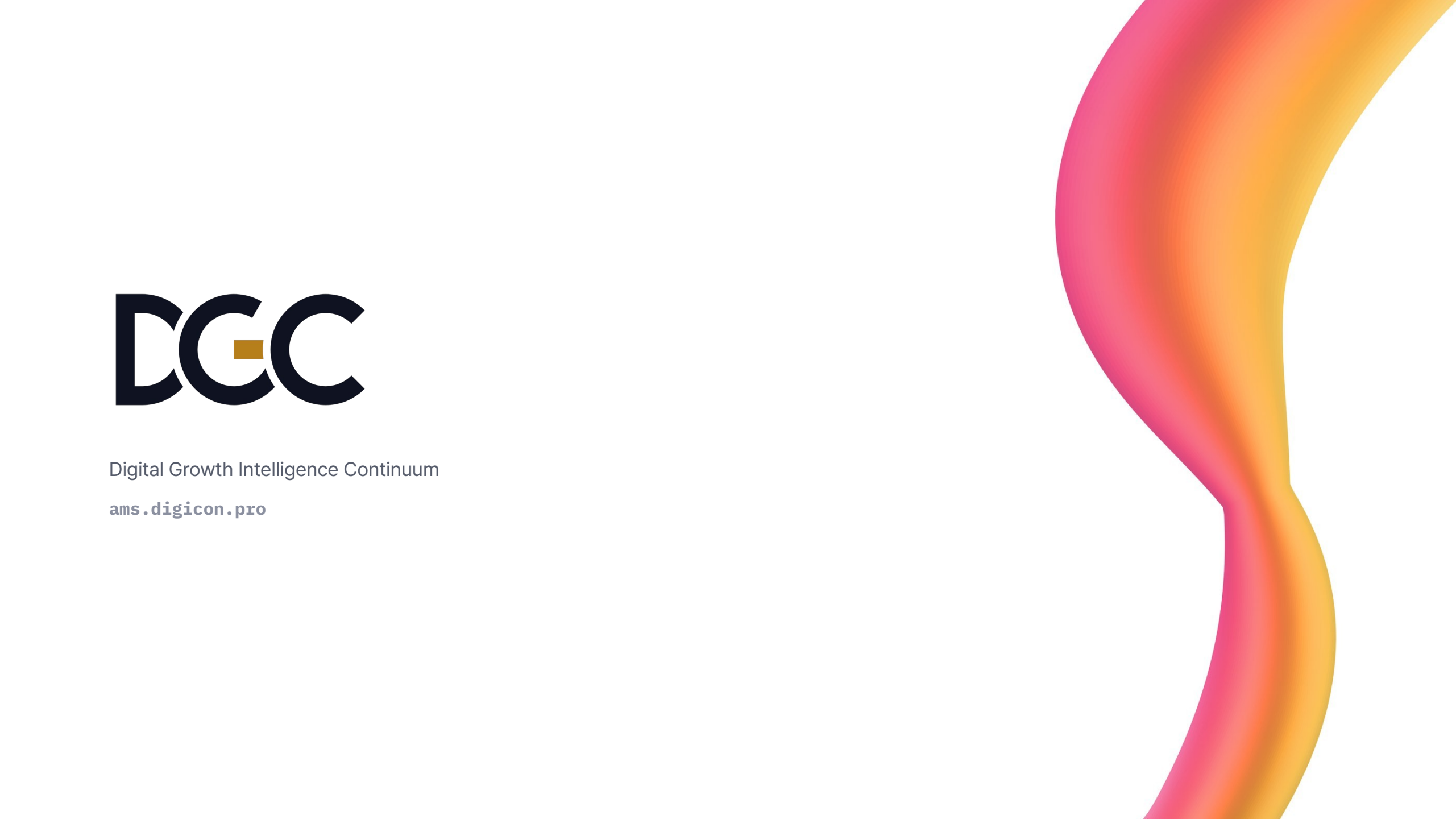
문의

Jay Jin, 대표

jay@digicon.pro · ams.digicon.pro

Digicon Holdings Pte. Ltd. · Singapore

결과는 **고객사만 보유**하고, 디지콘은 **공개하지 않습니다**.



Digital Growth Intelligence Continuum

ams.digicon.pro