

상표출원의 거절이유로 적용할 수 있는 유튜브 채널에 대한 기준 도출 연구

2026. 6.

제 출 문

본 보고서를 “상표출원의 거절이유로 적용할 수 있는 유튜브 채널에 대한 기준 도출” 수시과제의 최종보고서 제출합니다.

2026년 6월

○ 주관연구기관명: 한국지식재산연구원

○ 연 구 기 간 : 2026. 5. 11. ~ 2026. 6. 10.

○ 연구책임자 : 박윤석 (한국지식재산연구원 부연구위원)

[목 차]

제1장 서론	3
제1절 연구의 배경	3
제2절 연구의 목적	3
제2장 연구방법	4
I. 정량적 연구	4
II. 정성적 연구	4
제3장 유튜브 채널명 상표등록의 쟁점	4
I. 상표법 제34조 제1항 제13호 분석	4
1. 조문 분석	4
2. “국내 또는 외국의 수요자에게 인식된 요건” 해석	5
3. 제13호와 제9호(주지성)와 제11호(저명성)의 구별	5
II. 인식도 판단의 기준	6
1. 판례·심사기준상 인식도 판단 요소	6
2. 유튜브 구독자 1만 기준의 도입	6
3. 유튜브 채널명의 정량적 인지도 측정 가능성	6
제4장 연구 설계	8
I. 연구 설계 개관	8
II. 유튜브 채널 표본 수집	8
1. 수집 절차	8
2. 키워드 표본과 무작위 표본의 구별	10
(1) 키워드 표본	10
(2) 무작위 표본	10
3. 채널별 수집 항목 및 자동생성 채널 제외	11
III. 외부 언급량 수집(네이버 검색API)	12
1. 네이버 뉴스 블로그 카페 수집 설계	12

2. 채널명의 검색 쿼리 변환 문제	12
3. AI 기반 검색 쿼리 생성	14
4. 언급량 노이즈 검증 및 보정	14
제5장 연구 결과	16
제1절 정량적 분석	16
1. 수집된 채널의 기초통계량	16
2. 구독자 구간별 외부 언급 발생률	17
3. 노이즈 검증 및 인지도 보정 결과	17
4. 인지 발생 임계점(S자 곡선)	19
(1) 키워드 검색 채널 표본	19
(2) 무작위 검색 채널 표본	19
5. 인지 가속 변곡점	21
제2절 정성적 분석	23
제6장 연구 결과 분석	25

제1장 서론

제1절 연구의 배경

- 유튜브 채널이 브랜드 수익 기반으로 성장하면서 시청자 또는 소비자들에게 인지도를 얻은 채널명을 제3자가 상표로 선점 출원하는 사례가 나타나고 있음
 - 다수의 구독자를 보유한 국내 또는 해외의 유튜브 채널이 증가하고 있어, 유튜브 채널 명칭과 동일·유사한 출원상표에 대하여 적용하는 현행 심사 기준의 개선이 필요
- 채널 운영자 자신의 채널명을 보호받거나 타인의 무단 선점을 저지하려면 해당 채널명이 수요자에게도 인식된 표지인지가 쟁점이 될 수 있음

※ 상표법 제34조 제13호

13. 국내 또는 외국의 수요자들에게 특정인의 상품을 표시하는 것이라고 인식되어 있는 상표(지리적 표시는 제외한다)와 동일·유사한 상표로서 부당한 이익을 얻으려 하거나 그 특정인에게 손해를 입히려는 등 부정한 목적으로 사용하는 상표

- 현재 상표 심사 실무에서 유튜브 채널의 인지도 판단 기준으로 구독자 1만명을 참고하고 있으나 이 기준이 SNS가 활발하게 이루어지는 현재의 인터넷 환경에서도 적절한 기준인지 확인이 필요함
- 또한 유튜브 채널명에 대한 38류 상표등록에 대한 적절한 심사 기준을 확인할 필요가 있음

제2절 연구의 목적

- 위와 같은 연구 배경에 따라 본 연구는 다음의 핵심 질문을 확인하는 것을 목적으로 함

- | |
|---|
| <ul style="list-style-type: none">○ (문제제기 1) 구독자 수가 어느 수준일 때 채널이 외부 매체에서 실제로 언급·인식되기 시작하는가?○ (문제제기 2) 구독자수가 어느정도 되어야 유튜브 채널명이 시청자들에게 인지되었다고 볼수 있는 통계적으로 타당한 임계점은 어디인가? |
|---|

제2장 연구방법

I. 정량적 연구

- 본 연구에서는 유튜브가 제공하는 YouTube Data API v3 공개 데이터에서 분야별 유튜브 채널을 수집하고 해당 유튜브 채널이 네이버 뉴스, 블로그, 카페에 언급되는 양을 추적 조사하여 이용자들에게 얼마나 언급이 되고 있는지를 파악함

II. 정성적 연구

- 정량적 연구에서 얻은 데이터를 기반으로 유튜브 업계에서 활동하고 있는 전문가 5인을 섭외하여 전문가적 소견에서 유튜브 채널 인지도에 대한 의견을 듣고 정량적 연구 결과에 대해서 전문가들의 의견을 비교 분석함

제3장 유튜브 채널명 상표등록의 쟁점

I. 상표법 제34조 제1항 제13호 분석

1. 조문 분석

- (조문 내용) 제34조 제1항 제13호는 "국내 또는 외국의 수요자들에게 특정인의 상품을 표시하는 것이라고 인식되어 있는 상표(지리적 표시는 제외)와 동일·유사한 상표로서, 부당한 이익을 얻으려 하거나 그 특정인에게 손해를 입히려고 하는 등 부정한 목적으로 사용하는 상표"의 등록을 거절하는 규정임
 - ※ 제14호 관계에 대해서 선사용표장이 '지리적 표시'인 경우를 규율하며, 표장의 종류가 다를 뿐 제13호와 해석상 차이가 없음
 - ※ (요건) 성립 요건은 세 가지로 ① 선사용상표가 국내 또는 외국 수요자에게 '특정인의 상품을 표시하는 것'으로 인식되어 있을 것, ② 출원상표가 그 선사용상표와 동일·유사할 것, ③ 출원인에게 부정한 목적이 있을 것
- (입법 연혁) 1997. 8. 22. 법률 제5355호로 구 상표법 제7조 제1항 제12호로 신설(1998. 3. 1. 시행)된 후, 2016. 2. 29. 법률 제14033호 전부개정으로 제34조

제1항 제13호로 이전됨. 해외여행·시장개방·인터넷 보급에 따라 선출원주의를 악용해 타인의 신용·명성이 체화된 상표를 모방 출원하는 폐해를 방지하려는 취지임

- (규정 취지) 대법원은 본호의 취지에 관해, 국내외 수요자에게 특정인의 상품 표시로 인식된 상표가 국내에 등록되어 있지 않음을 기화로 제3자가 이를 모방한 상표를 등록·사용하여 권리자의 국내 영업을 방해하거나 손해를 가하거나 부당한 이익을 얻을 목적으로 사용하는 상표는 그 등록을 허용하지 않는다는 데 있다고 판시함 (대법원 2011. 7. 14. 선고 2010후2322 판결 등)

2. “국내 또는 외국의 수요자에게 인식된 요건” 해석

- (인식도 완화) 신설 당시에는 모방 대상 상표가 '현저하게 인식되어 있는 상표(주지상표)'일 것을 요구하여 모방상표 차단에 한계가 있었으나, 2007. 1. 3. 법률 제 8190호 개정으로 "현저하게 인식되어 있는"을 "인식되어 있는"으로 완화하여 인식도 요건을 낮춤
- (요구 정도) 따라서 현행법상 선사용상표는 주지상표에 이를 필요가 없고, 국내 또는 외국의 수요자 사이에 '특정인의 상품을 표시하는 정도'로 인식되어 있으면 충분함. 이 정도는 제12호 후단의 수요자기만 상표 요건과 동일하게, 국내 일반거래에서 그 상표·상품이라 하면 곧 특정인의 상표·상품으로 인식될 수 있는 정도를 말함 (대법원 2011. 7. 14. 선고 2010후2322 판결)
- (특정인의 의미) 여기서 '특정인'이란, 수요자나 거래자가 그 상표를 사용하는 자가 누구인지 구체적으로 인식하지는 못하더라도 익명의 존재로서 당해 상품의 출처를 인식할 수 있는 경우를 포함함
- (지리적 범위) 현행법은 "국내 또는 외국의 수요자"로 규정하므로, 외국의 특정 국가에서만 특정인의 상품 표시로 인식되어 있는 상표도 비교 대상 선사용상표가 될 수 있음.
- (판단 기준시) 판단 기준시는 제34조제2항에 따라 상표등록출원을 한 때를 기준으로 한다는 명시 규정을 둠

3. 제13호와 제9호(주지성)와 제11호(저명성)의 구별

- (인식 정도 차이) 제9호는 주지상표, 제11호는 저명상표를 보호하며, 이들은 제13호보다 높은 인식의 정도(주지·저명)를 요구함

- **(제13호의 위치)** 제13호는 주지·저명에 이르지 못하고 '특정인의 상품 표시로 인식되는 정도'이면 적용되는 대신, '부정한 목적'이라는 주관적 요건을 추가로 요구함
 - 대법원은 'BCBG'로 구성된 선사용상표가 제13호·제12호 후단의 '특정인의 상품'으로 인식될 정도에는 이르렀으나 제9호의 '주지성'은 취득하지 못하였다고 판시한 바 있음 (대법원 2011. 7. 14. 선고 2010후2322 판결)

II. 인식도 판단의 기준

1. 판례·심사기준상 인식도 판단 요소

- **(인식도 판단)** 선사용상표가 특정인의 상품 표시로 인식되었는지는 그 상표의 사용기간·방법·태양 및 이용범위, 거래실정 또는 사회통념상 객관적으로 널리 알려졌는지 여부 등을 기준으로 판단함 (대법원 2010. 7. 15. 선고 2010후807 판결 등)
- **(부정한 목적)** '부정한 목적' 유무는 특정인 상표의 인식도, 창작성의 정도, 표장의 동일·유사성, 출원인과 특정인 사이의 교섭 유무와 그 내용, 당사자의 관계, 출원인이 등록상표를 이용한 사업을 구체적으로 준비하였는지 여부, 상품의 동일·유사성 내지 경제적 건련관계, 거래실정 등을 종합적으로 고려하여 판단함 (대법원 2011. 7. 14. 선고 2010후2322 판결 등)
- **(전형 사례)** 전형적으로 부정한 목적이 인정되는 사례로는, 외국에서 알려진 타인의 상표가 국내에 미등록임을 기회로 이를 선등록한 후 고액으로 되팔거나, 국내 대리점 계약을 강제하거나, 외국 권리자의 국내 시장 진입을 저지하기 위해 출원하는 경우 등이 있음

2. 유튜브 구독자 1만 기준의 도입

- **(적용)** 유튜브 채널명에 제13호를 적용하려면, 그 채널명이 수요자에게 '특정인의 상품(채널·콘텐츠)을 표시하는 것'으로 인식되는 정도에 이르렀는지를 판단하여야 함
- **(구독자 수 원용)** 그러나 전통적 인식도 판단요소(사용기간·광고·매출 등)는 개인 크리에이터 채널에 그대로 적용하기 어려워, 측정이 간편한 '구독자 수'를 정량 지표로 원용

- **(검증 필요성)** 다만 구독자 수라는 단일 정량 기준이 '특정인의 상품 표시로 인식되는 정도'와 정합하는지는 실증적 검증 필요

3. 유튜브 채널명의 정량적 인지도 측정 가능성

- **(구독자 수의 지표로서의 한계)** 누적치 한계로서 구독자 수는 구독의 누적치로서, 현재 시점에 채널이 외부에서 실제로 인지·회자되는 정도를 직접 반영하지 못함
- **(인식과의 괴리)** 구독자가 많아도 외부 매체에서 거의 언급되지 않는 채널이 존재하여, 구독자 수와 외부 인지 사이에는 괴리가 있음. 제13호의 인식은 '수요자에게 특정인의 상품 표시로 인식'되는 것인데, 구독자 수만으로는 인지의 발현 여부를 포착하기 어려움
- **(외부 매체 언급량을 보조 지표 필요성)** 자발적 언급으로 뉴스·블로그·카페 등 제3자 매체의 자발적 언급은, 채널이 수요자에게 인식되었음을 보여주는 객관적 정황 자료가 됨
 - 이는 부정한 목적 판단요소 중 하나인 '특정인 상표의 인식도'를 뒷받침하는 객관적 자료로서 기능할 수 있음

제4장 연구 설계

I. 연구 설계 개관

※ 일반적으로 인지도가 높으면 인터넷 검색에서 노출이 많이 되는 것은 경험적 사실이기 때문에 유튜브 채널이 인지도가 높을수록 포털에서 노출이 많아진다는 것을 전제함

- YouTube Data API로 채널 표본을 수집하고, 각 채널명을 네이버에서 검색해 외부 언급량을 측정한 뒤, AI 검증으로 노이즈를 보정하고, 구독자-인지도 관계의 임계점을 통계적으로 추정하는 설계
- ※ YouTube Data API(v3)는 구글이 공식 제공하는 공개 데이터 조회용 인터페이스로, 유튜브 채널·영상·재생목록의 공개 정보를 사람이 화면에서 일일이 확인하는 대신 정해진 요청으로 받아 구조화된 데이터(JSON) 형태로 돌려주고 공개 통계만 다루며 비공개·개인정보는 포함되지 않음
- ※ 일일 호출량은 쿼터(quota) 단위로 제한되어(키별 일 10,000 units 등) 유튜브 채널 수집에 시간적 비용적 한계가 존재하였음
- ※ 본 연구에서 사용된 3가지 기능
 - ① 검색(search.list): 키워드·무작위 문자열로 채널을 찾아 채널 식별자(channel_id)를 확보
 - ② 통계 조회(channels.list): 확보한 식별자로 구독자 수·총 조회수·총 영상 수·개설일·국가 설정 등을 일괄 조회
 - ③ 업로드 조회(playlistItems.list): 채널의 업로드 목록에서 가장 최근 업로드 시각을 조회(활성 여부 판정용)

II. 유튜브 채널 표본 수집

1. 수집 절차

- (수집 절차) YouTube Data API v3를 통해 수집하였으며, 검색(search.list)으로 채널 식별자(channel_id)를 확보하는 1단계, 확보한 식별자로 채널 통계를 일괄 조회(channels.list)하는 2단계, 각 채널의 마지막 업로드 시각을 조회(playlistItems.list)하는 3단계로 구성하였음

- (검색 노출 기준) 키워드 표본과 무작위 표본의 채널 검색(search.list) 단계 모두에 지역 코드(regionCode=KR)와 관련 언어(relevanceLanguage=ko)를 적용하였음. 다만 이 파라미터는 한국에서 노출·검색되는 채널을 우선하도록 정렬하는 기준일 뿐 채널 국적을 한정하는 필터가 아니므로, 수집된 표본에는 외국 국적·외국어 채널도 포함됨
- ※ (문제: 수집의 편향성) 키워드 표본은 연구자가 14개 분야·130개 키워드를 직접 골라 검색한 것이고 어떤 키워드를 선택하고 그 키워드로 어떤 채널이 상위 노출되는가에 따라 표본이 특정 분야·인기 채널 쪽으로 쏠릴 수 있음(선택 편향). 즉 키워드 표본만으로는 유튜브 채널 전체 모습을 대표한다고 보장하기 어려움
- ※ (무작위 표본의 역할) 의미 없는 무작위 문자열로 검색하면 사람의 키워드 선택 의도가 개입하지 않아, 분야·인기와 무관하게 채널이 추출됨. 키워드 선택이라는 인위적 개입이 없었다면 모집단이 어떤 모습인가를 보여주는 대조군(control group) 역할
- ※ 실제로 무작위 표본은 글로벌·비한국어 채널을 더 많이 포함하여(표 3-2) 임계·변곡점이 더 높게 나옴
- (국적 무관 설계) 채널 국적을 한정하지 않은 것은 의도된 설계임. 제13호는 "국내 또는 외국의 수요자"에게 인식된 상표를 보호하므로, 한국에서 검색·노출되는 채널이라면 국적과 무관하게 분석 대상으로 보는 것이 정합

<국가 키워드>

구분	키워드 표본	무작위 표본
한국 (KR)	47.8%	13.9%
외국	8.4%	18.9%
미설정	43.8%	67.2%

2. 키워드 표본과 무작위 표본의 구별

(1) 키워드 표본

- **(분야 설계)** 유튜브 콘텐츠 전반을 포괄하는 14개 분야(뷰티/패션, 게임, 음악/댄스, 취미, 엔터테인먼트, 영화/애니메이션, IT/기술/과학, 라이프스타일, 여행/아웃도어, 푸드/쿠킹, 스포츠/건강, 뉴스/정치/이슈, 교육/강의, 경제/금융/재테크)를 설정
 - 각 분야를 대표하는 키워드(분야당 7~13개, 총 130개)로 채널을 검색하였음.
 - 분야를 구분한 것은 특정 장르에 표본이 쏠리지 않고 생태계 전반을 고르게 포착하기 위함임 (분야별 키워드 전체 목록은 부록 A 참조)
- **(검색 풀 ①)** 관련도 순(order=relevance)으로 키워드당 최대 500개 채널을 수집하여, 해당 주제에서 인기·관련도가 높은 상위 채널을 확보하였음
- **(검색 풀 ②)** 개설일 최신 순(order=date)에 최근 5년 이내 개설 조건을 결합하여 키워드당 최대 200개 채널을 추가 수집함으로써, 풀①이 놓치기 쉬운 신생·소형 채널을 보강하였음
- **(누적·중복 제거)** 검색 API는 1회 호출당 최대 50건을 반환하므로, 페이지 토큰을 이용한 연속 호출로 목표 수량까지 누적하였음. 동일 채널이 여러 키워드·풀에서 중복 검색될 수 있어 channel_id 기준으로 중복을 제거하여 고유 채널만 편입하였으며, 그 결과 키워드 표본은 원시 53,052개 채널로 구성되었음
 - ※ 수집할 때 channel_id를 기준으로 중복 채널을 제거하여 키워드 채널과 무작위 채널 내에서는 중복된 채널이 존재하지 않음. 따라서 키워드 채널과 무작위 채널 수집에 대해서 각각 통계적 분석시에는 중복이 발생하지 않음
 - ※ 다만, 키워드 채널과 무작위 채널 사이에는 255개(3.9%) 채널 중복이 존재하는데 이는 별도 수집절차를 통한 것으로 오히려 자연스러운 현상임. 그리고 통합적 분석시에는 중복을 제거하고 분석함

(2) 무작위 표본

- **(구성 취지)** 키워드 기반 수집이 내포하는 분야·키워드 선택의 표집 편향을 보정하기 위한 대조군으로, 의미 없는 무작위 문자열을 검색어로 사용하는 무작위 표본(Random Prefix/Keyword Sampling)을 별도 구성하였음

- **(검색어 생성)** 무작위 검색어는 한글 두 음절 조합 50%, 한글 자모 3개 조합 20%, 영문 알파벳 3글자 30%의 비율로 생성하였고, 난수 시드를 고정하여 동일한 검색어 집합이 재현되도록 하였음
- **(페이지 제한)** 무작위 검색어는 특정 주제로 수렴하지 않으므로 각 검색어당 1개 페이지(최대 50개 채널)만 수집하였음. 이는 무작위 검색어에서 페이지를 더 넘기는 것의 표본 대표성 기여가 적은 반면 쿼터 소모는 크다는 점을 고려한 설계임
- **(독립성)** 키워드 표본과 독립적인 추출 과정을 거쳐 원시 6,548개 채널로 구성되었으며, 두 표본 간 중복은 분석 단계에서 channel_id로 별도 확인하였음

3. 채널별 수집 항목 및 자동생성 채널 제외

- **(통계 수집)** 확보한 channel_id에 대해 channels.list를 50개 단위로 묶어 호출하여, 구독자 수·총 조회수·총 영상 수·채널 개설일·국가 설정값·구독자 수 공개 여부·업로드 재생목록 식별자를 수집하였음
- **(업로드 시각·결측)** 각 채널의 업로드 재생목록에서 playlistItems.list로 가장 최근 업로드 시각을 조회하여 활성 여부 판정에 사용하였고, 구독자 수를 비공개로 설정한 채널은 해당 값을 결측으로 표시하여 구분하였음
- **(자동생성 제외)** YouTube Music이 음원에 대해 자동 생성하는 'Topic' 채널(채널명이'○○○ - Topic' 형식)은 운영 주체가 없는 자동 생성물이므로 인지도 분석 대상에서 제외하였음
- **(쿼터 관리)** API 호출은 키별 일일 쿼터(검색100 units/호출, 통계·업로드 조회1 unit/호출, 키당 일 한도10,000 units)에 제약되므로, 다중API 키를 순환 사용하여 가용 쿼터를 확장하고, 쿼터 소진 시 직전까지의 결과를 보존한 채 중단하여 다음 수집 회차에서 이어받도록 하였음

4. 활성 채널의 정의와 필터링

- **(최근 1년 이내 업로드 기준)** 최근 1년 이내 영상을 올린 채널만 분석 대상으로 하며, 휴면 채널은 현재 인지도 측정이 어려워 제외함

<최종 수집 유튜브 채널 표본>

표본	수집 방식	원시 채널	분석 대상(활성·fit)
키워드 표본	14개 분야 키워드130개 검색	53,052	15,124
무작위 표본	무의미 음절·자모 검색(대조군)	6,548	1,738

III. 외부 언급량 수집(네이버 검색API)

1. 네이버 뉴스 블로그 카페 수집 설계

- **(측정 개념)** 채널의 외부 인지도를 제3자 매체가 해당 채널을 자발적으로 언급한 정도로 조작적 정의하였으며, 측정 도구로 국내 최대 검색 포털인 네이버의 검색 API를 사용하여 채널명에 대한 매체별 검색 결과 건수를 수집하였음
- **(매체 구분)** 외부 언급의 성격이 매체에 따라 다르므로 뉴스·블로그·카페 세 매체를 구분하여 수집하였음. 뉴스는 공식 보도 차원의 인지, 블로그는 개인의 자발적 소개·리뷰, 카페는 커뮤니티 내 회자를 각각 반영함
- **(수집 단위)** 채널마다 세 매체에 대해 검색API를 호출하여 각 매체의 총 검색 결과 건수(total)를 수집하고, 세 매체 건수를 합산한 총 언급량(total_mentions)과 매체별 건수를 함께 보존하였음
- **(0건과 실패 구분)** '검색결과 0건'과 '호출 실패'를 명확히 구분하였음. 0건은 그 자체로 '외부 언급이 없다'는 유의미한 인지도 정보인 반면, 호출 실패(요청 한도 초과·시간 초과 등)는 단순 수집 누락이므로, 각 호출 결과를 상태 코드(정상/요청한도초과/시간초과 등)로 기록하여 분석에서 혼동되지 않도록 하였음
※ 호출실패시 재시도 하도록 코딩하였음
- **(한도·체크포인트)** 검색API의 키별 일일 한도(약24,000회)에 맞추어, 일정 채널마다 중간 결과를 저장하는 체크포인트 방식을 적용하여 한도 도달 시 중단 후 다음 회차에서 미수집 채널만 이어받도록 하였음

2. 채널명의 검색 쿼리 변환 문제

- (문제 ①) 한·영 병기채널명이 한글과 영문을 함께 표기하는 경우(예: '애니살롱 ANISALON'), 영문이 한글의 음차(로마자 표기)이거나 같은 의미의 번역이면 영문을 검색어에 포함시킬 때 무관한 영문 문서가 다수 검색되어 노이즈가 됨
 - (문제 ②) 채널명이 보통명사인 경우(예: '저', '사람', '이유')에는 일상적인 글에서 해당 단어가 무수히 사용되어, 채널과 무관한 문서가 수억 건 단위로 검색되어 인지도 지표를 심각하게 왜곡함
 - (보완방법 ①) 정확매칭으로 검색어 양쪽에 쌍따옴표를 부여하는 정확매칭을 적용하여 해당 단어 배열이 정확히 포함된문서만 검색되도록 하였음
 - ※ (보완방법 ①은 채널명 정제) 한·영 병기 채널명에서 영문 음차·번역부를 제거하고 한글 핵심부만 검색어로 사용함. 유튜브 채널명이 수집될 때 원본 채널명은 한글과영어가 병합된 경우가 많음. 이 경우 영어까지 검색하면 검색결과가 0으로 나올 가능성이 매우 높아짐. 따라서 영어 제거
 - '애니살롱 ANISALON' → '애니살롱'
 - '짤캠핑 ZZan Camping' → '짤캠핑'
 - '파이스토리(PIESTORY)' → 파이스토리
- 그러나 영어를 삭제하면 안되는 경우도 존재
- 채널명이 'AI식당', '이상엽Math' 이런 경우에는 영어를 삭제하지 않고 전채로 채널명을 검색해야 해서 이러한 경우의 수를 LLM을 통해 정제함
- (보완방법 ②) 등급 분류정확 매칭만으로는 보통명사 오탐을 제거하지 못하므로, 채널명을 검색 적합도에 따라 네 등급(fit/moderate/risky/unfit)으로 사전 분류하여 분석 단계에서 fit한 채널만 분석함
 - ※ 정확매칭만으로도 '저'·'이유' 같은 보통명사는 무관 문서가 여전히 대량 매칭되므로, 채널명을 검색 적합도에 따라 4등급(fit 고유명·신뢰도 높음/ moderate / risky 일반어 단독·오탐 위험 큼 / unfit 검색 부적합)으로 사전 분류하여, 주 분석은 fit에 해당하는 채널명만 네이버 수집에 활용하여 분석함
 - ※ 'fit'은 고유성이 높아 검색 신뢰도가 높은 채널명, 'risky'는 일반어 단독 등 오탐 가능성이 큰 채널명에 해당함
 - fit 예시 : “재테크의 여왕 슈엔슈” , “주식투자훈련소”

- moderate 예시 : “코코메이크업” , “10분영어”
- risky 예시 : “창문” , “단화”

3. AI 기반 검색 쿼리 생성

- **(활용 배경)** 채널명에서 식별성높은 검색어를 추출하는 작업은 단순 규칙만으로 한계가 있어 대규모 언어모델(Claude)을 보조적으로 활용하였음
- **(생성 방식)** 모델은 각 채널명에서 한글 핵심부를 추출하고 음차에 해당하는 영문을 제거하되, 방송사·브랜드 약어 등 식별자 역할의 영문은 보존하는 방식으로 검색 쿼리를 생성하였음
- **(검색 쿼리 생성 기준 및 보완)** 다수 채널을 묶어 배치로 처리하였고, 인공지능이 검색 쿼리를 생성할 때 동일한 입력에 동일한 출력이 나오도록 제한하였고 어떤 이유이든 인공지능이 검색 쿼리 생성에 실패한 경우 보완책으로 규칙 기반(정규표현식) 방식으로 대체하여 누락이 없도록 하였음. 이렇게 생성된 정제 쿼리를 정확매칭과 결합하여 최종 검색에 사용하였음

4. 언급량 노이즈 검증 및 보정

- **(노이즈 문제의 발생 원인)** 일반어 채널명은 무관한 글이 대량 매칭됨(예: 채널명 "저"→ 수억 건). 원시 건수는 인지도로 직접 쓸 수 없음
- **(AI 표본 판별을 통한 진짜 언급 비율 추정)** 표본 채널 1,999개의 검색 결과 94,414건을 AI(Claude)가 직접 읽어 진짜/가짜를 판별
- **(머신러닝 모델을 통한 전체 채널 인지도 보정)** 인공지능의 판별 결과를 정답으로 삼아 유튜브 채널에 대한 실제적인 언급 비율을 예측하는 모델을 구축하고 이 2000개 채널을 기준으로 머신러닝을 진행 한 후 이 모델을 이용하여 네이버 검색량을 추정하여 채널 보정 인지도(adj_total) 산출
 - ※ 네이버의 검색결과가 단어만 매칭되면 검색결과로 알려주기 때문에 실제 유튜브 채널명이 검색결과에서 언급된 것인지 확인 필요
 - ※ 네이버에서 구한 검색결과 전체가 수천만건에 달하는데 이것을 모두 확인해서 실제로 채널명이 언급된 것인지 확인하기에는 시간과 비용의 한계
 - ※ 따라서 3단계 통계 추정 방식을 이용함

- ① 전체 검색결과에서 표본(1,253개 채널·검색결과 94,414건)만 AI가 직접 읽어 각 결과의 진짜/가짜를 판별
 - ② 그 판별 결과를 정답(label)으로, 채널 특성(이름 등급·길이·구독자·언급량 등)으로 '진짜 언급 비율(정밀도)'을 예측하는 머신러닝 모델을 만들어 학습
 - ③ 학습된 모델로 표본에 없는 전체 채널의 정밀도를 예측하고, 원시 검색 건수에 곱하여 보정 인지도(adj_total)를 산출
- ※ (가장 쉬운 예시) 네이버 검색결과가 1200개가 나왔다고 가정한다면 이중에서 100개를 순서대로 인공지능이 읽고 실제 채널명이 언급된 사례가 10개라면 10%이므로 1200×0.1 로 해서 120개로 산출하는 방식임. 그러나 네이버 검색결과 노출이 관련도 순서이므로 뒤에 나오는 검색결과는 거의 해당사항이 없을 가능성이 높아 이런 사항을 반영하여 예측함

제5장 연구 결과

제1절 정량적 분석

1. 수집된 채널의 기초통계량

○ (분포 특성) 구독자·외부 언급 모두 극심한 우편향

- 평균이 중앙값을 크게 상회함 (소수 대형 채널이 평균을 끌어올림). 예컨대 키워드 표본의 구독자 평균은 약10.8만이나 중앙값은 693에 불과함
- 수집된 채널의 85%정도가 구독자수 만명 이하임
- 아래 그래프의 외부언급량은 네이버, 블로그, 카페에서 언급된 건수를 합한 것임

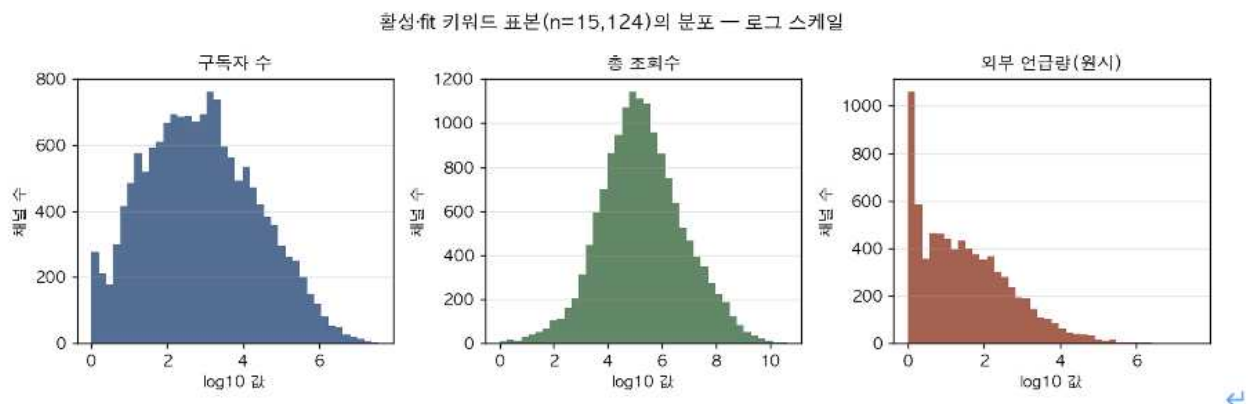


그림 4-1. 활성·fit 키워드 표본(n=15,124) 구독자·조회수·언급량 분포 (로그 스케일)

- ##### ○ (표본별 기술통계 비교) 기준 모집단으로 아래 기술통계는 이후의 회귀·임계 분석과 동일한 분석 대상, 즉 활성·fit 채널(키워드15,124 / 무작위 1,738)을 기준으로 산출

통계량	키워드 (n=15,124)	무작위 (n=1,738)
구독자 평균(명)	108,419	203,864
구독자 중앙값(명)	693	702
구독자 95%(명)	289,000	749,150
총 외부언급 평균(건)	5,564 (보정전)	3,170 (보정전)
총 외부언급 중앙값(건)	1	0
총 외부언급 95 분위(건)	1,447	548
영상 수 중앙값(개)	90	103

표 4-1. 활성·fit 채널 기초 통계량 (외부 언급은 검증·보정 전 원시 건수, 분석 대상과 동일 모집단)

2. 구독자 구간별 외부 언급 발생률

- (단조 증가) 구독자 구간이 높을수록 '외부 언급 1건 이상' 확률 $P(\geq 1)$ 이 단조 증가하며, 키워드 표본은 구독자 1천~3천 명 구간(51.5%)에서 처음으로 50%를 넘어섬

※ 언급량 분포가 극심한 우편향(소수 채널이 매우 큼)이라 평균은 소수 대형 채널에 끌려 왜곡되므로, 중앙값이 구간의 대표값으로 더 적절함

구독자 구간	n	$P(\geq 1)$	$P(\geq 10)$	$P(\geq 100)$	총언급 중앙값
[0, 100)	4,487	35.5%	21.3%	10.9%	0
[300, 1K)	1,869	46.6%	27.6%	16.1%	0
[1K, 3K)	1,854	51.5%	27.8%	14.3%	1
[3K, 10K)	1,499	59.8%	31.3%	14.5%	1
[10K, 30K)	1,224	66.8%	37.1%	16.9%	3
[30K, 100K)	1,009	70.3%	41.9%	18.8%	5
[100K, 300K)	675	72.3%	51.7%	28.4%	13
[300K, ∞)	741	74.1%	56.8%	38.9%	26

표 4-2. 활성·fit 키워드 표본($n=15,124$) 구간별 외부 언급 발생 확률

3. 노이즈 검증 및 인지도 보정 결과

- (판별 결과) 94,414건 판별 결과 원시 검색 결과의 약 74%가 노이즈로 확인
 - 실제 언급은 26.0%.

판정	비율
실제 채널 언급 (yes)	26.0%
우연 매칭·무관 (no)	72.4%
모호 (unclear)	1.6%

표 4-3. AI 사후 검증 판별 결과 (총 94,414건)

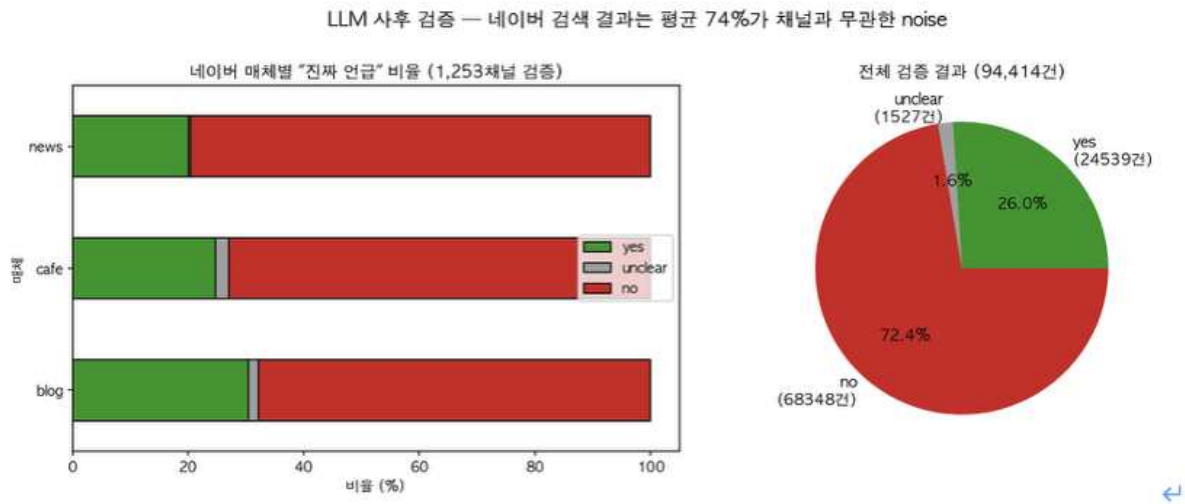


그림 4-3. AI 검증 판별 분포

○ (보정 전후 언급량 변화)

데이터셋	원본 평균	보정 평균	제거된 노이즈
키워드 fit	5,564	1,028	약 81%
무작위 fit	3,170	517	약 84%

표 4-4. 원본 대비 보정 외부 언급량 (모델 R^2 0.25~0.32)

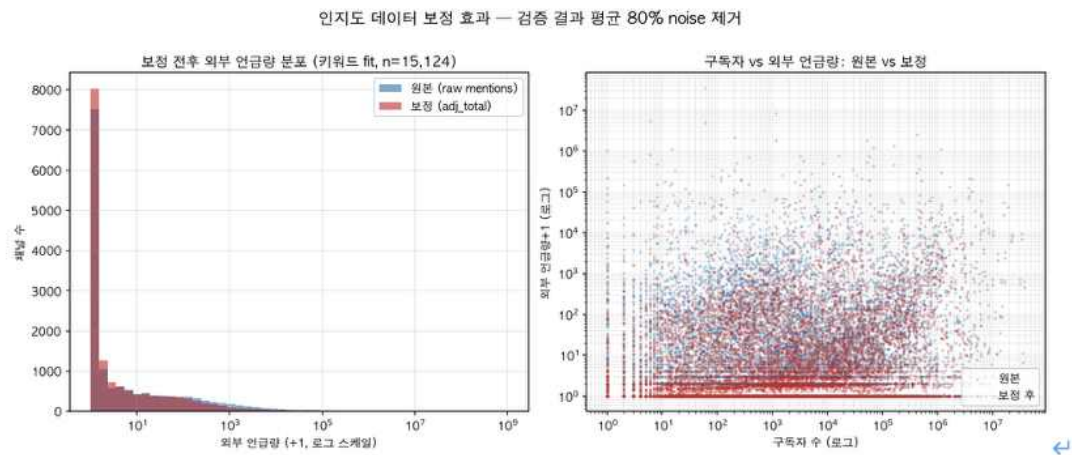


그림 4-4. 원본 vs 보정 외부 언급량

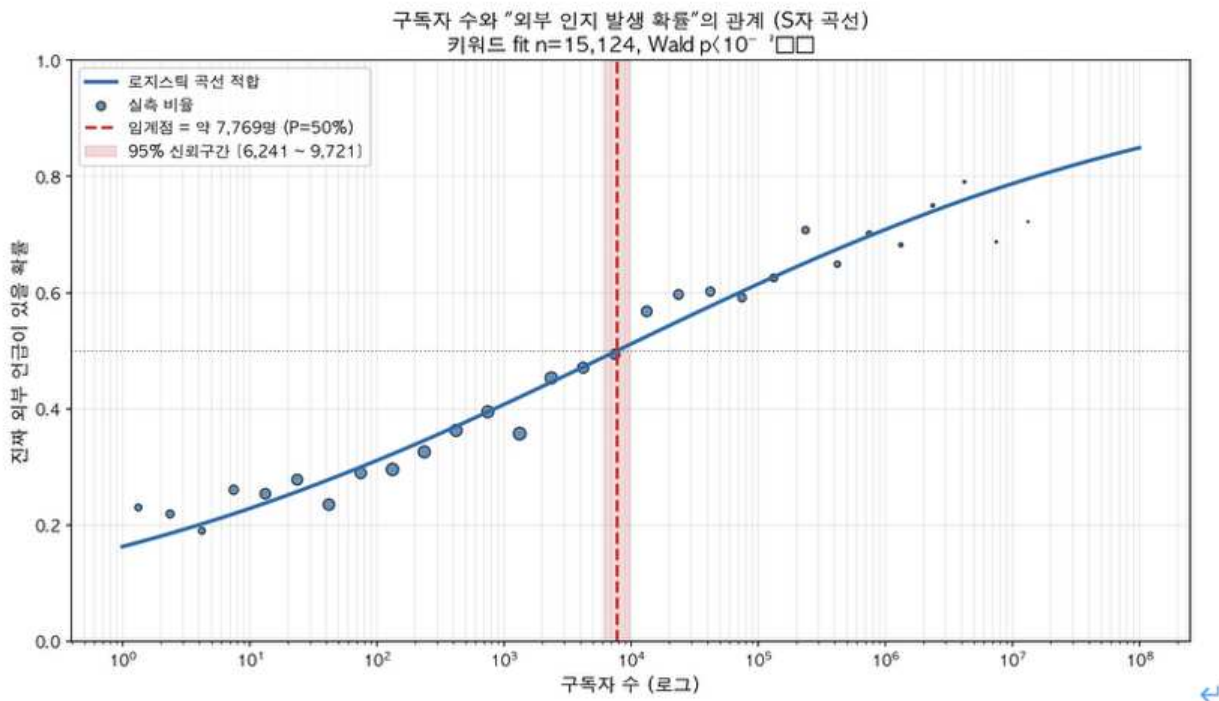
4. 인지 발생 임계점(S자 곡선)

(1) 키워드 검색 채널 표본

○ (외부 언급 발생 확률) 50%를 통과하는 지점 약 7,800명

항목	값
인지 발생 임계 (P=50% 통과)	약 7,800명
95% 신뢰구간	6,241 ~ 9,721
로지스틱 적합 유의성	$p < 10^{-200}$

표 4-5. 인지 발생 임계 (키워드 fit, $n=15,124$)



(2) 무작위 검색 채널 표본¹⁾

○ (무작위 표본) 무작위 표본에 같은 방식(4파라미터 로지스틱, 활성·fit, 보정 언급)을 적용하면, 인지 발생 임계는 약 148,000명으로 키워드 표본보다 현저히 높음

1) 추출된 표본에서 외부언급량이 매우 적어(전체 언급률이 14.7%) 따라서 이 148,000명의 값은 신뢰성이 낮음. 따라서 변곡점 그래프로 판단해야함

- 무작위 표본이 글로벌·비한국어 채널을 더 많이 포함하여(표 3-2 참조) 동일 구독자 대비 국내 언급이 늦게 발생하기 때문임

항목	값
인지 발생 임계 (P=50% 통과)	약 148,000명
표본 전체 언급률	14.7%
적합 방식	4파라미터 로지스틱(보정 언급)

표 4-6. 무작위 표본 인지 발생 임계 (활성-fit, n=1,738)

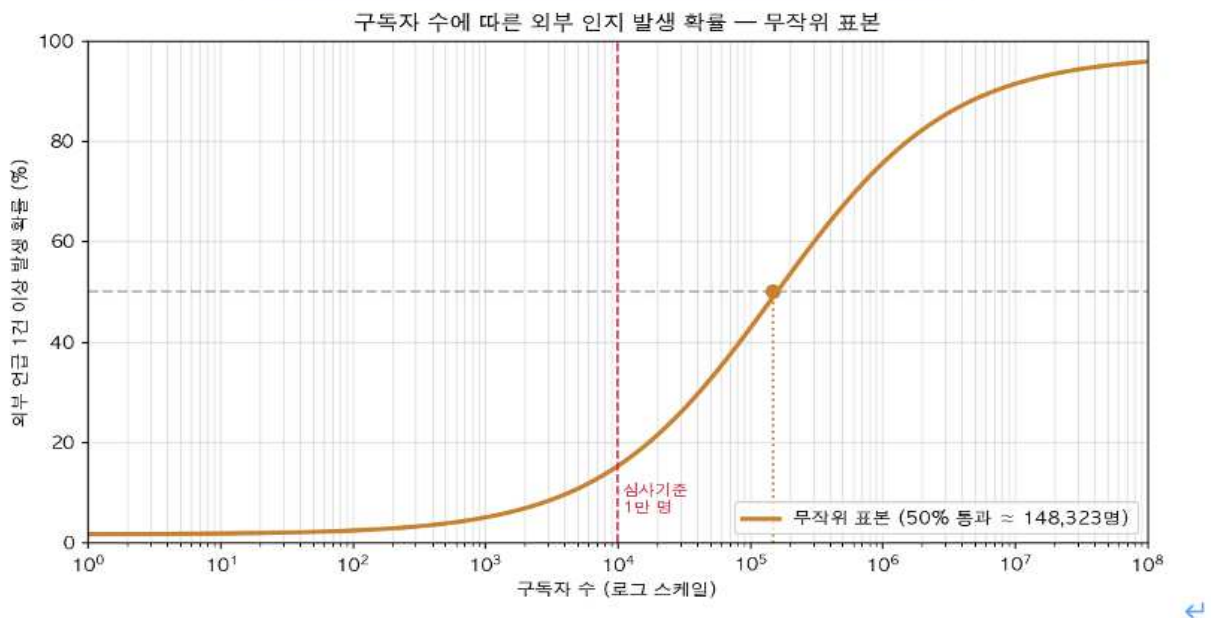


그림 4-6. 무작위 표본 구독자 vs 외부 인지 발생 확률 (S자 곡선)

- ※ 두 표본의 S자 임계(키워드 약 7,800 / 무작위 약148,000)는 표본 전체 언급률 차이에 민감하여 절대 격차가 과장될 수 있으므로, 표본 간 비교는 보다 안정적인 변곡점을 기준으로 함
- ※ S곡선의 50% 임계(곡선의 중점)는 그 표본 전체에서 ‘외부 언급이 있는 채널의 비율’(전체 언급률) 수준에 강하게 좌우됨. 전체 언급률이 낮으면 곡선이 전반적으로 낮게 깔려 50%에 도달하는 데 훨씬 많은 구독자가 필요함
- ※ 우리가 수집한 키워드 채널과 무작위 채널중 무작위 채널의 경우 해외 채널도 다수 포함되어 있어 구독자수가 높아도 네이버에 언급될 확률이 낮아서 키워드 채널에 비하여 구독자 수가 훨씬 많아야 네이버에 언급될 가능성이 높아짐

- ※ 따라서 키워드 표본과 무작위 표본 두 개를 비교하기는 적절치않다는 의미임
- ※ 결국 키워드 표본은 언급이 비교적 흔하고, 무작위 표본은 매우 드물어(전체 언급률 14.7%) 곡선이 낮게 깔려서 그 결과 임계가 무작위 약 148,000명 vs 키워드 약 7,800명처럼 절대 격차가 과장되어 보인다는 의미임
- ※ 쉽게 생각하면 해외 유튜브 채널은 보통의 한국 유튜브 채널보다 훨씬 인지도가 높아야 네이버에서 검색된다고 생각하면 됨

5. 인지 가속 변곡점

- (구독자-인지도) 회귀의 기울기가 약 9,200명 지점에서 0.12 → 0.37로 약 3배 가속

데이터셋·조건	표본	변곡점(구독자)
키워드 fit (1년, 보정)	15,124	9,230
무작위 fit (1년, 보정)	1,738	36,300

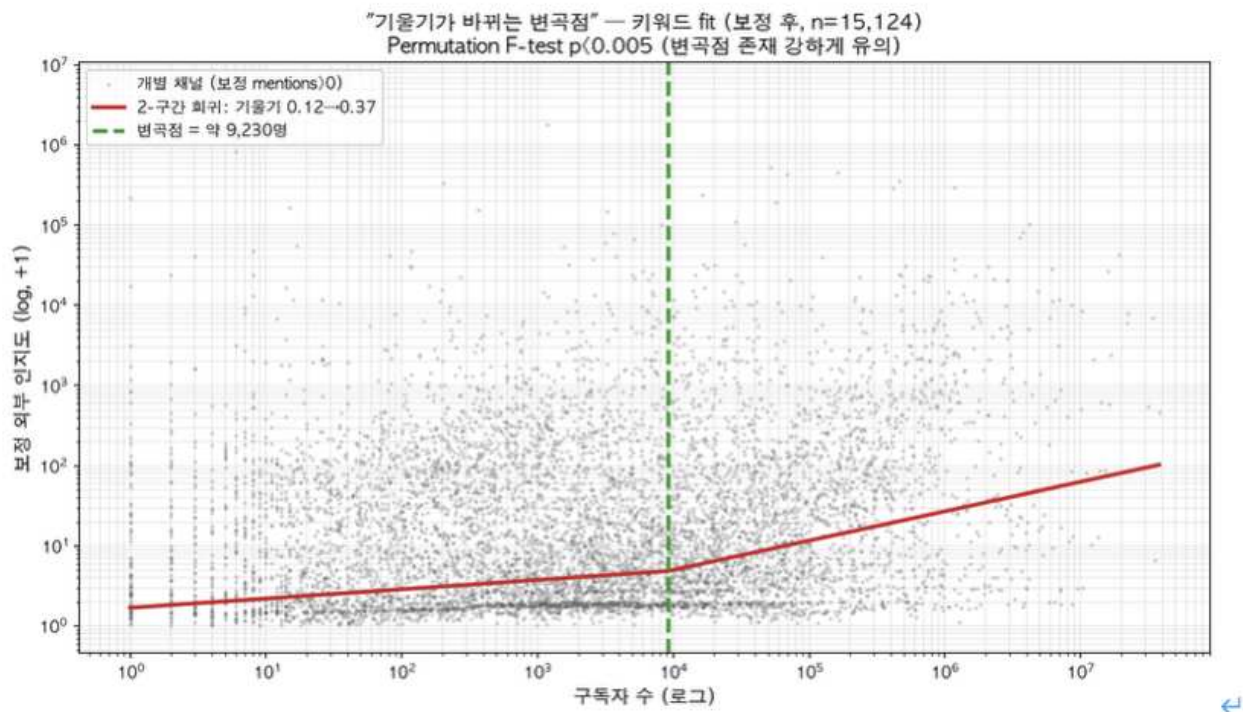


그림 4-7. 구독자 vs 인지도 변곡점 (구간별 회귀)

- (변곡점 비교) 인지 가속 변곡점은 키워드 표본 약 9,230명, 무작위 표본 약 36,300명으로 약 3.9배의 차이를 보였으며, 이는 표본 구성이 아니라 채널의 언어·시장 특성 차이에서 비롯된 것으로 해석됨
- ※ 언어·시장 특성이란 채널이 주로 어떤 언어로, 어떤 국가의 시청자를 대상으로 하는가를 뜻함. 한국어·국내 대상 채널은 국내 인지와 직접 연결되고, 외국어·해외 대상 채널은 국내 인지와 괴리가 존재함(해외에서 유명해도 국내에서 유명하지 않을 수 있고 그 반대도 가능함)
- ※ 외부 언급량은 국내 최대 포털인 네이버에서 측정한 것이어서 무작위 표본은 한국에서 검색·노출되더라도 글로벌·비한국어 채널을 더 많이 포함하므로, 구독자가 많아도 네이버 언급은 늦게·적게 발생함
- ※ 따라서 무작위 표본의 변곡점이 더 높은 구독자에서 나타나는 것은 표집 방식이 아니라, 채널의 언어권·주 시장이 '네이버 기반 국내 인지'와 정합하는 정도가 다르기 때문으로 해석된다는 의미임

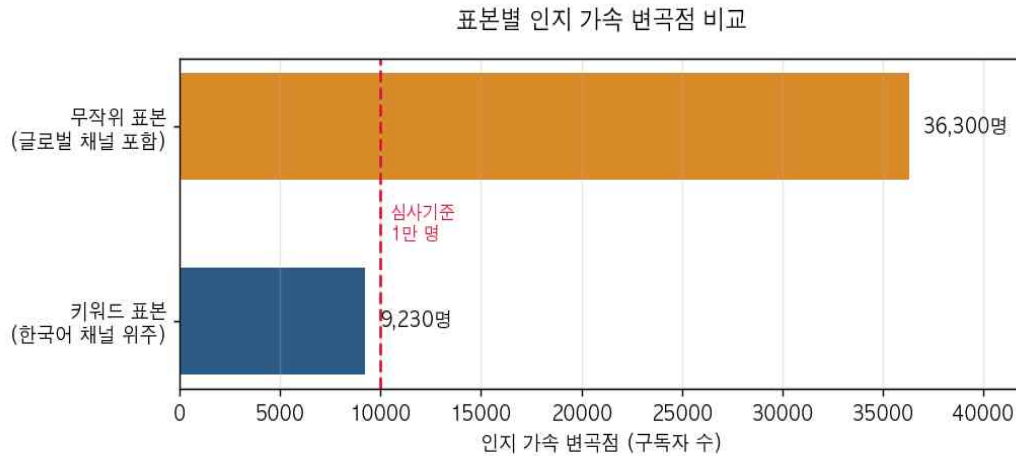


그림 4-11. 표본별 인지 가속 변곡점 비교

6. 정량적 분석에 대한 소결

- 정량적 분석의 전제조건인 가정은 유튜브 채널 이름이 시청자 및 소비자들에게 인식되어 있다면 네이버와 같은 포털에서 언급될 것이라는 점
- 이 가정을 상표법 제34조 13호와 비교하여 보면 포털에 언급이 될수록 채널명이 타인의 채널로 인식될 확률이 높아진다는 의미
- 정량적 연구결과 임계점 분석에서 키워드 표본에서는 구독자수가 7,800명 정도 되어야 네이버 포털에서 언급될 확률이 50%를 넘어가고 무작위 키워드 표본(해외채널 포함)의 경우 구독자 수가 14만명 정도 되어야 네이버 포털에 언급될 확률이 50%가 넘는 것으로 나왔는데 이 14만이라는 구독자 수는 유명한 해외 채널이라도 우리나라 포털에 언급될 확률이 한국 채널보다 낮기 때문에 구독자수 격차가 크게 나타난 것으로 해석됨

※ 따라서 수치가 과장된 것일수 있어 가속도 분석으로 다시 해봄

- 네이버 채널 언급과 관련하여 변곡점 분석으로 해본 결과 키워드 표본의 경우 구독자수가 9,230명이 되면 네이버 포털에 언급되는 분량이 많아지기 시작하고 무작위 키워드 표본의 경우 구독자 수가 36,300명이 넘어갈 때 네이버 포털에 언급량이 많아지기 시작한다고 판단됨
- 이러한 점을 종합해 보면 유튜브 채널의 구독자 수가 만명에서 3만명 사이의 구간에 있는 경우 네이버 언급량이 증가하는 속도가 빨라지게 된다고 판단됨
- 해석상 유튜브 채널의 구독자수가 만명에서 3만명 사이인 경우 상표법 제34조 제13호에서 의미하는 타인의 채널로 인식될 수 있는 시작점이 될 수 있다는 것을 의미함

제2절 정성적 분석

- (개요) 유튜브 관계자들과 FGI를 통해 정량적 분석 결과에 대한 전문가 의견을 수집하고 설문지에 대해서 회의 개최
 - 총 5회에 걸쳐서 SBS, CJ ENM, 트레저헌터, SLL 등 SNS 시장에서 영향력을 가지고 있는 기업들의 유튜브 담당자와 인터뷰 진행
- (주요 결과)
 - (구독자수 만명에 대하여) 유튜브 채널 구독자수 10,000명은 달성하기 어려운 조건이라는 점에는 모든 전문가가 동의함
 - ※ 다만, 채널명의 인지도와 채널운영자의 인지도가 유사하게 가기 때문에 연예인과 같은 공인이 유튜브 채널을 운영하는 경우 구독자수가 5만명 정도는 되어야 이제 채널이 사람들에게 알려졌다고 보는 의견 있음 (근거:광고주가 대부분 구독자수 5만명을 원하는 경우가 많음)
 - ※ 개별의견으로 현재의 만명 기준보다 높은 수준이 필요하다고 예를들어 10만명 정도가 필요하다는 의견이 있었음 (근거 : 만명은 성장 가능성이 있는 채널이고 10만명이 업계나 언론에 영향력있는 채널로 보기 시작하는 경우가 많다고 생각함)
 - ※ 유튜브에서 인정하는 최소기준인 실버버튼이 구독자수 10만임
 - ※ 경우에 따라 어뷰징(거짓 회원수 늘리기)이 발생할수도 있어 투자만 하면 구독자수 5만명까지는 쉽게 늘리수 있다고 함
 - ※ 영상 시청률·평균 조회수, 구독자 대비 조회수 비율, 일일 접속자수, 참여율을 종합적으로 비교해야 외부 인지도를 확인할 수 있음
 - (분야별로 구독자수와 인지도는 다름) 언어가 필요한 채널과 언어가 필요없는 ASMR, 먹방 채널 경우 채널 공급도 많고 시청자도 많아 영상시청에 언어가 중요한 역할을 하는 채널보다 언어가 필요없는 영상을 공급하는 채널의 구독자수가 높아야 인지도가 있다고 판단할 수 있음
 - ※ 정치, 시사 채널보다 커버댄스, 먹방 채널이 구독자수가 높아야 인지도가 있다고 볼 수 있음
 - ※ 음악처럼 거의 모든 사람이 즐기는 보편적 분야는 모수가 커서 인지되려면 구독자수도 많아야 함. 반대로 특정 악기·마이너 장르 전문 채널은 향유 인구가

적어 구독자수가 적어도 그 안에서는 충분히 알려진 채널로 통합.

- (유튜브로 수익창출은 만명보다 작아도 가능) 유튜브를 통한 수익창출은 여러 가지 방법이 있어서 구독자수 만명이 안되어도 수익창출은 가능함
 - ※ 유튜브 뿐만 아니라 인스타그램, 틱톡 등 여러 계정을 운영하고 있어 유튜브만 판단하는 것보다 종합적인 2차적 이용 상황을 SNS로 판단하는 것도 필요함
- (유튜브 채널과 그 영상의 2차적 사용이 중요) 유튜브 채널이 유명한 경우 그 채널의 영상이 2차적 이용(다른 네트워크에서 바이럴되거나 이용되는 현상)이 되는 정도로 시청자 및 소비자들의 인지도를 파악하는게 중요함

제6장 연구 결과 분석

□ 핵심 문제제기에 대한 정량적 분석 결과

(문제제기 1) 구독자 수가 어느 수준일 때 채널이 외부 매체에서 실제로 언급·인식되기 시작하는가?
(문제제기 2) 구독자수가 어느정도 되어야 유튜브 채널명이 시청자들에게 인지되었다고 볼 수 있는 통계적으로 타당한 임계점은 어디인가?

- (문제제기 1 결과) 구독자 수가 9,000명~30,000명 정도가 되면 유튜브 채널명이 외부에 최초로 언급되기 시작하면서 인지도가 상승하는 시점이라고 판단됨
- (문제제기 2 결과) 구독자 수가 30,000명 정도가 되면 유튜브 채널명이 외부에서 인지되어 사람들이 특정인의 채널로 인식될수 있다고 판단됨
 - ※ 상표법 제34조 제1항 제13호의 판단 지표로 활용이 가능하다고 판단 됨

□ 상표법 제34조 제13호와 관련된 시사점

- (우리나라 제34조 제13호의 구조) 국내외 수요자들에게 특정인의 상품을 표시하는 것이라고 인식되어 있는 상표가 존재하고 부정한 목적이 존재해야함
 - 선사용상표의 인식정도는 과거 "현저하게 인식되어 있는"을 “인식되어 있는”으로 선사용상표의 인식도의 정도를 완화하는 방향으로 상표법 개정이 이루어짐
 - 현행 상표법에서는 선사용상표의 인식정도를 주지상표라고 해석할 수는 없고, 국내

또는 외국의 수요자 사이에 특정인의 상품을 표시하는 정도로만 인식되어 있으면
충분(상표법 주해, 754면)

※ 상표 인식의 정도 순서

저명(상표법 제34조 제1항 제11호) > 주지(상표법 제34조 제1항 제9호) > 사
용에 의한 식별력(상표법 제33조 제2항) > **특정인 인식(상표법 제34조 제1항
제13호)**

○ (우리나라 제34조 제13호에 대한 판결 분석) 대법원 2014. 2. 13. 선고 2013후2460 판
결 등

- 모방대상상표가 국내 또는 외국의 수요자 사이에 특정인의 상표로 인식되어 있는
지는 상표의 사용기간, 방법, 태양 및 이용범위 등과 거래실정 또는 사회통념상 객
관적으로 상당한 정도로 알려졌는지 등을 기준으로 판단하여야 함(대법원 2019.
8. 14. 선고 2017후752 판결)
- 익명 출처로 족함. 수요자가 출처의 구체적 명칭을 알 필요 없이 "동일한 하나의
출처"로 인식하면 충분함

○ (네이버 언급과 특정인의 상표 인식 관계) 네이버에서 특정인의 유튜브 채널명이
언급된 경우 유튜브 채널명을 특정인 또는 특정 채널에서 제공하는 영상의 출처로
인식할 수 있게됨

- 유튜브 채널명이 채널명으로서 유명해진 것과 운영자의 인지도로 인해 유명해 진
것은 구별할 실익이 있지만²⁾ 유튜브 채널의 주 목적이 특정 콘텐츠를 제공하는
것인 점을 고려하면 유튜브 채널명 자체가 포털에 언급된 것은 유튜브 서비스의
표지로서 언급된 것으로 볼 수 있음
- 다른 규정 (9호, 12호, 제33조 제2항)과 제13호의 구조에서도 비교한 것처럼 제
13호에 해당되는 식별력은 상표법상 요구되는 식별력 중에서 가장 낮은 단계의
식별력임
- 이러한 점을 고려하면 네이버 포털의 언급 빈도수도 식별력 판단의 보조자료로 이
용될 수 있다고 판단 됨

2) 예를들어 유튜브 채널 운영자가 사회적 이슈(불법, 정치 등의 사회적 쟁점)에 집중을 받는 경우는 유튜브 채널의 인
지도와 관계없이 운영자의 인지도만 높아지는 경우가 가능할 것으로 생각됨. 이런 경우 채널의 구독자수와 유튜브
운영자의 인지도 간에 격차가 발생할 수 있음

□ 연구 결과와 특정인의 상표 인식 관계

○ (유튜브 채널 구독자와 네이버 언급 결과)

- 유튜브 채널의 구독자 수가 많을수록 유튜브 채널명이 네이버 포털에서 언급되는 횟수가 많아짐. 이것은 네이버 포털에 유튜브 채널명이 언급될수록 유튜브 채널이 특정인의 출처로 인식되고 있다는 점을 확인할 수 있음(이러한 점은 연구의 경험적 가정조건이 있음)
- 정량적 분석 결과 한국어로만 구성된 채널명의 표본에서는 구독자 수가 다음과 같음

데이터셋·조건	변곡점(구독자)
키워드 표본	9,230
무작위 표본	36,300

- 변곡점이 의미하는 것은 유튜브 채널명의 네이버 언급량이 이전보다 갑자기 많아지는 구간을 의미하고 이것은 유튜브 채널명이 네이버에서 많이 언급되는 시점을 의미함
- 따라서 이 구간을 제13호에서 이야기 하는 유튜브 채널명이 특정인의 채널이라고 인식되는 구간이라고 판단할 수 있는 지표 중 하나에 해당될수 있다고 보임
- 정성적인 분석결과에서도 전문가들의 의견도 최소 만명에서 5만명의 구독자를 보유해야 특정인의 유튜브 채널이라고 인식할 수 있다고 함

□ 연구 결과에 대한 연구자 의견

○ (연구자 의견) 구독자수 만명은 특정인의 채널 인지에 대한 종점이 아니라 시작점이라고 판단됨

- 통계적 사항과 전문가 의견에서 외부에 인지된 채널이 되기 위해서는 구독자수가 최소한 3만명 이상은 되어야 한다고 판단됨
- 그리고 구독자수 뿐만 아니라 종합적인 요소가 판단에 필요함

※ 종합적인 판단 기준 예시

- ① 영상 시청률·평균 조회수
- ② 구독자 대비 조회수 비율 - 좀비채널(활성화되어 있지 않은 채널) 판단용
- ③ 일일 접속자·참여율(댓글·좋아요)

④ 외부 매체 언급량(본 연구 지표)

⑤ 영상의 2차적 이용(타 플랫폼 인스타그램, 틱톡 등 바이럴)

○ (고려사항)

- 전문가의견에서 구독자수 단독으로 채널명의 인지도를 판단하기는 어렵다고 봄
- 전문가들은 구독자수가 만명에서 10만명 정도가 되어야 상표등록에서 보호받을 수 있는 특정 채널로 인식될 수 있다고 의견을 가짐
- 다만, 전문가 모두 실질적인 구독자 10,000명을 달성하는 것은 매우 어렵다고 판단함

붙임 2

분야별 검색 키워드

○ 14개 분야 키워드목록 (v1 수집 기준, 총 130개)

분야	키워드 (개수)
뷰티/패션	메이크업, 패션, 코디, 헤어스타일링, 네일아트, 스킨케어, 명품리뷰, 빈티지패션, 옷장정리 (9)
게임	게임방송, 롤, 마인크래프트, 오버워치, 모바일게임, 발로란트, 닌텐도스위치, 인디게임 (8)
음악/댄스	노래, 커버, K-POP, 보컬, 발라드, 힙합, 인디밴드, 트로트, 재즈, 클래식음악, EDM, 뮤지컬, 댄스커버 (13)
취미	캠핑, DIY, 만들기, 낚시, 사진편집, 디지털드로잉, 자전거타기, 가드닝, 드론촬영 (9)
엔터테인먼트	예능, 개그, ASMR, 토크쇼, 챌린지, 슬라임, 리액션, 몰래카메라, 시트콤, 패러디 (10)
영화/애니메이션	영화리뷰, 애니, 단편영화, 영화추천, 만화추천, 웹툰, 일본애니, 영화해석, 시리즈리뷰 (9)
IT/기술/과학	IT, 코딩, 과학, 개발자, 인공지능, 천체관측, 가젯리뷰, 컴퓨터조립, 생물학 (9)
라이프스타일	브이로그, 일상, 자취, 직장인브이로그, 인테리어, 육아, 1인가구, 미니멀라이프, 정리수납 (9)
여행/아웃도어	여행, 해외여행, 트레킹, 백패킹, 캠핑, 국내여행, 자전거여행, 호캉스, 글램핑, 일본여행 (10)
푸드/쿠킹	요리, 먹방, 베이킹, 한식레시피, 디저트레시피, 자취요리, 비건레시피, 반찬만들기 (8)
스포츠/건강	운동, 헬스, 스포츠, 홈트, 필라테스, 요가, 골프, 축구, 러닝 (9)
뉴스/정치/이슈	뉴스, 시사, 정치, 경제뉴스, 국제뉴스, 시사기획, 팩트체크 (7)
교육/강의	강의, 공부, 영어, 토익, 자격증, 수능, 일본어, 수능수학, 동기부여, 중국어, 영어회화, 자녀교육 (12)
경제/금융/재테크	재테크, 주식, 부동산, 비트코인, 스타트업, 가계부, 부업, 금융지식 (8)

○ 무작위 표본의 검색어는 고정된 목록이 아니라, 난수 시드(seed=42)로 알고리즘이 생성한 무의미 문자열임. 세 유형을 비율(한글 두 음절 50% / 한글 자모 3개 20% / 영문 3글자 30%)로 생성하며, 동일 시드에서 동일집합이 재현됨

※ 무작위 검색어는 사람이 손으로 고른 목록이 아니라 컴퓨터의 난수 생성기로 자

동 생성한 의미 없는 문자열임. 난수 생성기는 ‘시드(seed)’라는 시작 숫자를 입력으로 받아 정해진 계산으로 난수 열을 생성

※ 핵심은 시드가 같으면 난수 생성기가 매번 정확히 같은 난수 열, 즉 같은 검색어 집합을 만들어낸다는 것임. 본 연구는 코드를 누가 언제 다시 실행해도 동일한 무작위 검색어 집합이 그대로 재현됨

유형	생성 방식	검색어 예시
한글 두 음절	초·중·종성 무작위 조합 2자	인낮, 쓴멤, 일갯, 둘챙, 짜넌, 푼칠
한글 자모 3개	자모 무작위 3개	ㄷ ㅅ ㅁ, ㅍ ㅈ ㅊ, ㅍ ㅈ ㅊ, ㅇ ㄴ ㅋ, ㅂ ㅋ ㅌ
영문 3글자	알파벳 무작위 3자	gsw, kgu, ehx, mic, sml, rri, oei