

H_k 구성하기

- 정보 가리기 및 호기심 유발
- 과도한 감정적 형용사 및 선정적 표현
- 형식적 아티팩트 및 강조 기호

.가공 패턴 레지스트리 (Pattern Registry)

각 패턴은 가이드라인에서 정의한 낚시성 유형을 기반으로 하며, 사람이 "낚시"라고 인지하는 대표 표현들을 포함합니다.

[P1] 정보 가리기 및 호기심 유발 (Curiosity Gap)

- 패턴 정의:** 핵심 정보를 대명사나 부사로 숨기거나 문장을 미완성 상태로 끝내어 클릭을 유도하는 방식입니다.
- 대표 표현 세트 H_1**
 - 결론 지연: 알고보니, 사실은, 결국, 반전
 - 이유/대상 강조: 이유는?, 어찌나, 누구?, 그 내막은
 - 미완성/지시어: ...보니, ...했더니, ...길래, 이것
 - 정규표현식(Regex): ? (물음표 사용), ... (말줄임표 사용), 어떻길래|무슨 일

[P2] 과도한 감정적 형용사 및 선정성 (Sensationalism)

- 패턴 정의:** 사건의 객관적 전달보다 독자의 즉각적인 감정 반응(경악, 공포 등)을 이끌어내는 자극적 단어를 사용하는 방식입니다
- 대표 표현 세트 H_2**
 - 경악/충격 계열: 충격, 경악, 발칵, 멘붕, 기겁
 - 자극적 동사: 폭로, 포착, 단독공개, 포착, 터졌다
 - 최상급/강조: 역대급, 소름, 깜짝, 숨막히는
 - 정규표현식(Regex): \W!\{1\} (느낌표 반복), 급상승|화제

[P3] 형식적 아티팩트 및 강조 기호 (Formatting Artifacts)

- 패턴 정의:** 내용의 본질과 무관하게 모델이 '낚시성' 혹은 '정식 기사'라고 오인하게 만드는 기호나 특정 문구의 반복적 배치입니다
+ 1
- 대표 표현 세트 H_3**
 - 특수 기호 강조: [속보], [단독], [종합], [공식]
 - 시스템성 문구: 1보, 재배포 금지, 무단 전재, 현장포착
 - 정규표현식(Regex): \W[\.*\W] (대괄호 패턴), 【.*】 (특수 괄호), [!W?]{2,} (문장부호 혼합 사용)

패턴 ID	패턴명	주요 표현 (Hk)
\$P_1\$	정보 가리기/호기심 유발	알고보니, 사실은, 결국, 반전, 이유는?, 어쩌나, ...보니, ...했더니, ...길래, 이것
\$P_2\$	감정적/선정적 표현	충격, 경악, 빨칵, 멘붕, 기겁, 폭로, 포착, 단독공개, 터졌다, 역대급, 소름
\$P_3\$	형식적 아티팩트/기호	[속보], [단독], [종합], [공식], 재배포 금지, 무단 전재, [!?, [], [] 등 특수기호]

패턴 ID	실제 데이터에서 매칭된 Hk 후보 (사람 정의)	발견 건수
P1	"알고보니", "결국", "이유는?", "사실은", "반전", "...", "?" 등 4	53건
P2	"충격", "경악", "폭로", "!", "단독공개", "소름" 등 5	2건

P3	"[속보]", "[단독]", "재배포 금지", "[[]]", "【】" 등 ⁶	17건
----	--	-----

패턴/카테고리	신규 표현 (Newk)	IG 점수	특징 및 정성 해석
경제 (DIRECT_EC)	"짓는다더니...", "오지고", "신박한", "쏟아진다..."	1.14 ~ 1.49	구어체 비속어나 유행어를 사용한 감성적 유입 유도 (P1/P2의 확장형)
기술 (DIRECT_ET)	"도플갱어", "비하인드", "사연은?", "테이트가"	1.10 ~ 1.18	연예/방송 콘텐츠에서 특정 키워드(도플갱어, 비하인드) 자체에 대한 편향
사회 (DIRECT_GB)	"잡힌18개월", "남아의", "박살나다", "죽어나나?"	1.50 ~ 2.45	자극적인 수치(18개월)나 극단적인 동사(박살나다)에 대한 민감도

정치 (DIRECT_PO)	"정유라", "법카", "윤석열35%", "체포동의안"	0.89 ~ 1.02	특정 인물이나 정치적 쟁점 단어 자체를 낚시성 지표로 오인 (Topic Bias)
생활 (DIRECT_LC)	"발가벗긴", "음악가", "살아남은", "어디길래?"	1.62 ~ 3.06	호기심 유발형 질문뿐 아니라 자극적인 형용사(발가벗긴)의 영향력 확인

패턴 유형	겹친 주요 단어/기호	분석 및 의미
P1 (호기심 유발)	"?", "알고보니", "사연은?", "어디길래?"	모델도 의문문과 결론 지연형 부사를 매우 강력한 낚시성 신호로 인지하고 있습니다.
P3 (형식적 기호)	"[", "]", "...", "!"	대괄호([])나 말줄임표(...) 같은 문장 부호가 기여도 상위권에 대거 포진했습니다. 모델이 기사의 형식적 틀을 중요하게 본다는 증거입니다.
P2 (선정성)	"박살나다", "역겹다", "터져..."	자극적인 동사나 형용사 일부가 모델의 상위 50개 리스트에 포함되었습니다.