

인공지능 기반 멀티모달 감성 불일치 분석을 통한 부상기술 성장 불확실성 분석: 대형언어모델 사례

권재현¹ · 신준석^{1,2*}

¹성균관대학교 기술경영전문대학원 / ²성균관대학교 시스템경영공학과

Analysis of Emerging Technology Growth Uncertainty through AI-empowered Multimodal Sentiment Incongruence Analysis: A Case Study of Large Language Models

Jae-Hyun Kwon¹ · Juneseuk Shin^{1,2}

¹Graduate School of Technology Management, Sungkyunkwan University

²Systems Management Engineering, Sungkyunkwan University

This study proposes a novel methodology and a quantitative metric, the Sentiment Incongruity Index (SII), for analyzing multimodal signals and their incongruence in video-based evidence on emerging technologies, particularly large language models (LLMs). Our approach integrates AI-based facial expression analysis and text mining into a unified multimodal framework to quantify discrepancies between textual sentiment and facially expressed sentiment. Applying this framework to LLMs as a case study reveals a consistent positivity bias in expert verbal expressions, and pronounced sentiment incongruence across both LLMs in general and their key sub-technologies. The findings further underscore the importance of experts' nonverbal cues for mitigating biases inherent in expert judgment and traditional text analytics, and lay the groundwork for incorporating nonverbal signals into emerging technology analysis.

Keywords: Multimodal Sentiment Analysis, Emerging Technology, Large Language Model, Sentiment Incongruence

1. 서 론

부상기술(emerging technology)을 조기에 파악하고 적시에 채택하는 역량은 기업 경쟁우위의 핵심요소이다(Teece, 2007). 나아가 기존 산업의 패러다임을 전환하고 새로운 사회경제적 가치를 창출하는 변화의 원동력이다(Zahra and Nielsen, 2002). 그러나 부상기술 중 일부만 지속적으로 성장하며, 다수는 정체가되거나 쇠퇴한다. 따라서 높은 성장 잠재력과 사회경제적 파급력을 지닌 부상기술을 파악하는 역량은 기업뿐만 아니라 산업과 국가 차원에서도 중요하다.

국내 학계 및 실무에서는 흔히 유망기술(promising technology)이라는 용어가 혼용되고 있으나, 유망이라는 표현은 미래의 성공 가능성에 대한 긍정적 편향을 내포할 수 있다. 이러한 이유로 본 연구는 가치 중립적이며 기술의 성장 과정에 필연적으로 수반되는 불확실성 등의 역동적 변화를 설명하는 부상기술이라는 용어를 사용했다. 부상기술은 기존 기술과의 차별성, 빠른 성장 속도, 관련 지식 및 기술 요소 간의 뚜렷한 응집성, 광범위한 사회경제적 영향력을 가지며, 높은 수준의 불확실성과 모호성을 지닌다(Rotolo *et al.*, 2015). 이러한 속성은 부상기술 측정에 있어 두 가지 문제를 야기한다. 첫째, 부상기술

* 연락저자 : 신준석 교수, 16419 경기도 수원시 장안구 서부로 2066 성균관대학교 기술경영전문대학원, Tel : 031-290-7607, Fax : 031-290-7610,
E-mail : jsshin@skku.edu

2026년 2월 4일 접수; 2026년 2월 24일 수정본 접수; 2026년 3월 6일 게재 확정.

의 사회경제적 영향력은 사후적으로만 확인할 수 있다. 둘째, 불확실성과 모호성은 정량화하기 어렵다. 결과적으로 미래 가치 창출 가능성이 높은 부상기술을 초기 단계에서 파악하고 평가하는 일은 여전히 도전적인 과제로 남아 있다(Rosenberg, 1996; Rotolo *et al.*, 2015).

전통적으로 부상기술 파악과 예측은 전문가 평가에 의존해 왔다. 전문가 평가 기반 기법들은 복수 전문가의 다양한 지식, 경험, 통찰력 활용이 장점이다. 그러나 연구자들은 개별 전문가의 주관적 판단이나 밴드왜건 효과(bandwagon effect)로 인한 편향을 공통된 문제로 지적해 왔다(Rowe and Wright, 1999). 또한 고도로 전문화된 특정 분야에 고착되어 급변하는 기술 환경에 유연하게 대응하지 못하는 한계가 있다(Landeta, 2006; Tetlock, 2005). 더욱이 융복합화로 기술의 복잡도가 증가하면서, 소수 특정 분야 전문가들의 제한된 지식으로 인한 평가 정확도 감소 문제도 심각해지고 있다.

과학기술문헌의 전자문서화와 컴퓨팅 파워의 비약적 발전은 대규모 문서 데이터 분석의 토대를 마련했다(Porter and Cunningham, 2005). 연구자들은 과학기술문서 데이터를 정량적으로 분석하는 계량서지학(bibliometrics)을 기술혁신 연구에 활용하기 시작했다. 보다 빠르고 정교한 문서 분석을 위해 자연어 처리(natural language processing, NLP) 기반 텍스트 마이닝(text mining) 기법이 도입되었다(Kyeibambe *et al.*, 2017). 연구자들은 기술 동향 분석, 신기술 발굴 등 여러 목적을 위해 서지학, NLP 기법, 전문가 평가를 복합한 다양한 과학기술문서 분석 방법을 제안했다(Érdi *et al.*, 2013).

부상기술 파악과 예측 분야에서도 기존 전문가 기반 기법들의 한계를 극복하기 위한 다양한 대규모 데이터 및 문서 분석 방법들이 제안되었다. 최근 연구들은 대형언어모델(large language model, LLM)과 인공지능(artificial intelligence, AI)을 활용해 기존 정량·정성적 방법론들의 통합을 시도하고 있다(Song *et al.*, 2023; Wang *et al.*, 2024). 특히, 텍스트, 이미지, 비디오 등 여러 유형의 데이터를 결합하는 멀티모달(multimodal) 분석은 기존 문서 기반 방법들의 한계를 극복할 수 있는 새로운 대안으로 주목받고 있다(Baltrušaitis *et al.*, 2018; Poria *et al.*, 2017).

문서 기반 부상기술 분석 및 예측 방법들에는 공통된 한계가 존재한다. 부상기술의 성장, 채택, 확산에는 기술의 우수성, 시장 수요, 사회적 수용성과 같은 다양한 요인들이 서로 상호작용하면서 영향을 미친다(Rogers, 2003). 기존 연구들은 이러한 상호작용이 항상 합리적이지는 않으며, 다양한 감성적·맥락적 요인들이 작용한다고 주장한다(Read *et al.*, 2011). 과학기술문서는 엄밀한 검증 과정을 거치면서 이러한 요인들을 소거한다. 따라서 결과물은 이성과 논리 중심이다. 또한, 특허를 포함한 신기술 관련 문서는 전문가들의 낙관적 편향(optimism bias)이나 전략적 오도(strategic misrepresentation)가 개입되어(Flyvbjerg, 2008) 기술의 긍정적 측면을 강조하고, 잠재적 위험은 생략하거나 축소하는 경향이 있다. 즉, 문서만으로는 부상기술의 개발, 채택, 확산과 관련된 이해당사자들의 감성적·

맥락적 진의를 파악하는 데 한계가 있으며, 긍정 편향된 결과를 얻을 확률이 높다.

빠르게 성장하는 영상 채널들은 위와 같은 문서 기반 분석의 단점들을 보완할 수 있는 기회를 제공한다. 기술과 제품에 대한 개발자, 투자자, 소비자의 견해들을 포함한 영상들은 다양한 미디어에서 급격히 증가하고 있다(Kaplan and Haenlein, 2010). 이러한 영상들은 문서와 같거나, 때로는 문서에 없는 정보를 제공한다. 또한, 최근 유튜브(YouTube)는 기존 미디어 플랫폼과 경쟁하며 전문적인 기술 정보와 시장 견해를 공유하는 핵심 채널로 부상하고 있다(Budzinski *et al.*, 2021). 나아가, 영상은 다양한 이해당사자들의 의견을 언어뿐 아니라 얼굴 표정, 목소리, 억양과 같은 감성적·맥락적 정보로도 제공한다(Ambady and Rosenthal, 1992; Budzinski and Gaenssle, 2020).

감성 불일치 가설(sentiment incongruity hypothesis)은 부상기술 분석을 위한 영상 데이터 활용에 이론적 기반을 제공한다. 전문가의 언어적 표현과 비언어적 표현 간에는 종종 불일치(incongruity)가 발생한다. 주요 원인 중 하나로 인지 부조화(cognitive dissonance)가 제시되어 왔다. 전문가가 언어적 견해와 다른 생각을 내적으로 가지고 있어서 심리적 불편함이 발생하고, 이러한 감성이 비언어적으로 표현된다는 것이다(Festinger, 1957). 특히 언어와 비언어적 신호가 불일치할 경우, 화자의 진의는 표정과 같은 비언어적 단서를 통해 전달될 수 있음이 보고되어 왔다(Mehrabian and Wiener, 1967). 예를 들어, 전문가가 부상기술의 성장 가능성에 대해 긍정적으로 말하면서도, 내적 불확실성이 높다고 생각할 경우, 얼굴 표정에 경직이나 불안이 발생할 확률이 높다는 것이다.

이러한 맥락에서 본 연구는 영상에 포함된 부상기술에 대한 언어적 표현과 화자의 표정에서 나타나는 감성 간 일치(congruity) 수준을 정량적으로 측정·분석하는 방법을 제시한다. 영상에 포함된 화자의 음성을 텍스트로 변환한 후, 감성 분석을 통해 부상기술에 대한 긍정·부정 수준을 측정하고, 동시에 화자의 얼굴 표정 감성을 분석한다. 최종적으로 텍스트와 표정 간 감성 일치 수준을 측정해 불일치 수준이 높은 부상기술을 파악한다. 다양한 영상에서 평균적으로 불일치 수준이 높게 나타나면, 전문가들이 언어적으로는 긍정적이지만 내적으로 불확실성과 모호성이 높다고 생각하는 부상기술일 확률이 높다.

본 연구에서 제안하는 방법은 세 가지 장점이 있다. 첫째, 기존 방법들의 한계였던 부상기술의 불확실성과 모호성을 정량적으로 측정할 수 있다. 둘째, 텍스트 분석의 긍정 편향을 멀티모달 분석을 통해 감소시킬 수 있다. 셋째, 소요시간이 많은 전문가 기반 방법들을 적용하기 어려운 급성장하는 부상기술에 대한 신속한 분석이 가능하다. 최근의 부상기술 중 LLM은 빠르게 성장하며, 불확실성이 높고, 텍스트상 긍정 편향이 높게 나타날 가능성이 큰 대표적 기술이다. 따라서, 본 연구의 방법론 적용에 적합하다. 영상 플랫폼에 해당 기술의 미래 전망에 대한 영상이 다양하고 풍부해 데이터 수집도 용이하다.

본 연구는 다음과 같이 구성된다. 제2장에서는 부상기술의

주요 특성과 불확실성을 고찰하고, 부상기술 파악 방법에 관한 선행연구를 살펴본다. 제3장에서는 데이터와 텍스트-영상 감성 불일치 분석 방법론에 대하여 설명한다. 제4장에서는 LLM에 대한 분석 결과를 정리한다. 제5장에서는 전문가 언어 표현의 긍정 편향과 감성 불일치 유형을 논의하며, 마지막으로 제6장에서는 본 연구에서 제시한 방법의 학술적 기여와 실무적 가치를 논하고, 주요 한계점과 향후 연구 방향을 제시한다.

2. 문헌 연구 및 이론적 배경

2.1 부상기술 특성과 불확실성

부상기술은 과학기술의 발전 과정에서 기존 패러다임을 위협하는 새로운 패러다임의 핵심요소로 간주되어 왔다(Dosi, 1982). 초기 연구자들은 시장구조 재편, 신산업 창출과 같은 부상기술의 사회경제적 영향력에 주목했다(Porter, 2002). 연구자들은 미래가치 창출의 원인을 부상기술의 속성에서 찾았고, 신규성(novelty)과 빠른 성장성(fast growth)에 주목했다(Martin, 1995). 그러나 두 가지만으로는 설명력이 부족했다. 후속 연구들은 기술 수렴(convergence)을 포함하는 속성인 응집성(coherence)을 추가하여, 부상기술을 과학기술 기반의 사회경제적 혁신 과정으로 정의했다(Day and Schoemaker, 2000). 연구자들은 세 가지 속성을 정량적으로 측정해 유망한 기술을 파악하는 다양한 지표와 방법론을 제시했다(Cozzens *et al.*, 2010).

그러나 위의 세 가지 속성이 모두 탁월한 부상기술도 사회경제적 가치 창출에 실패하는 경우가 많았다. 연구자들은 기술의 출현(emergence)과 성장 과정에서 나타나는 불확실성을 핵심 원인으로 제시했다(Boon *et al.*, 2008). 불확실성은 기술 발전이나 융합과 같은 기술 자체의 원인뿐 아니라, 정책이나 시장과 같은 외부환경의 변화에 의해서도 발생한다(Srinivasan, 2008). 즉, 언제 어떤 불확실성이 나타나 부상기술에 어느 정도의 영향을 미칠지 예측하기 어려운 것이다. 더욱이 불확실성의 원인, 유형, 영향력은 부상기술의 성장 과정에서 지속적으로 변화한다(Stirling, 2007). 따라서 동적 불확실성 관리는 훨씬 어려운 문제가 된다. 요약하자면, 불확실성의 예측, 파악, 관리는 부상기술 성장의 핵심이지만, 이를 성공적으로 실행하기는 쉽지 않다.

Rotolo *et al.*(2015)은 선행연구들의 논의를 종합해 부상기술의 특성을 급진적 신규성(radical novelty), 빠른 성장성(fast growth), 응집성(coherence), 현저한 영향력(prominent impact), 불확실성과 모호성(uncertainty and ambiguity)의 다섯 가지로 제안했다. 앞서 언급한 바와 같이, 성장성과 영향력은 초기 단계에서 파악이 어려운 사후적 특성이다. 따라서 성장 단계의 부상기술에는 신규성, 응집성, 불확실성과 모호성이 공존한다. 기존 연구들은 부상기술의 성장을 신규성을 바탕으로 응집성을 증가시키면서, 불확실성과 모호성에 대응하는 과정으

로 정의한다(An and Ahn, 2016). 성공적인 기업은 다양한 장애를 극복하고 부상기술의 고속성장과 사회경제적 영향력을 실현하며, 이를 통해 시장에서 경쟁우위와 지배적 위치를 선점한다(Lee *et al.*, 2019; Lieberman and Montgomery, 1988; Wiefels, 2002).

기술수용모델(technology acceptance model)은 일반적인 기술의 채택과 확산 관점에서 불확실성의 유사한 영향력을 지적한다. Davis(1989)는 사용자의 기술 채택은 유용성과 사용 용이성이라는 객관적 속성에 대한 주관적 인식에 의해 결정된다고 주장하였다. 그러나 후속 연구들은 두 가지 요인 외의 다양한 개인적, 거시적 요인들이 기술 확산에 미치는 영향력을 실증하였다(Venkatesh *et al.*, 2003). 특히 기술 수용 초기 단계에는 고객 관성, 정부 규제, 대체재 등 예상치 못한 요인들이 등장해 기술 확산을 정체시키는 경우들이 많았다. 즉, 기술수용 초기 단계에는 다양한 불확실성이 존재하며, 이를 극복한 후에야 비로소 기술은 초기 시장과 주류 시장 사이의 캐즘(chasm)을 넘어 본격적인 성장 궤도에 진입하게 된다(Moore, 1991; Wiefels, 2002). 나아가 기술 채택자의 빠른 증가를 위해서는 기존 채택자들의 긍정적 평가와 실용적 성과가 일정 수준 이상 누적되어야 한다. 심리학적 관점에서 이는 선행 채택자들의 행동과 평가가 후발 수용자들에게 강력한 사회적 증거로 작용하여 불확실성에 대한 심리적 장벽을 낮추기 때문이다(Cialdini, 2007). 이처럼 기술 확산은 혁신에 내재된 불확실성이 사회적 연결망을 통해 점진적으로 해소되어 가는 인지적 전환 과정으로 볼 수 있다(Rogers, 2003).

기존 연구들은 부상기술의 지속 성장이 일부 탁월한 속성만으로는 어렵다는 점을 강조한다. 신규성이 뛰어난 급진적 신기술의 사업화 실패 사례나, 빠르게 부상하던 기술의 오랜 성장 정체가 좋은 예이다(Christensen, 2015; Kirchberger and Pohl, 2016). 부상기술의 성장을 위해서는 신규성, 응집성과 같은 긍정적 속성의 극대화뿐만 아니라, 불확실성에 대한 대비가 필요하다. 전략 포사이트(strategic foresight) 연구들이 강조하는 바와 같이 불확실성 대비의 첫 단계는 측정이다(Vecchiato, 2012). 기존 불확실성 측정 방법은 정성적 전문가 평가와 서지학과 머신러닝(machine learning)을 복합한 텍스트 분석방법들이 주류를 이루고 있다(Tobback *et al.*, 2018). 주목할 점은 불확실성의 사회경제적 효과가 인간의 인지를 통해 발생한다는 것이다. 따라서 인지된 불확실성(perceived uncertainty)에 영향을 미치는 감성적, 심리적 요인들을 측정 방법에 반영할 필요가 있다. 그러나 아직까지는 부상기술과 관련된 감성적, 심리적 요소나 인지된 불확실성을 정량적으로 측정하는 방법에 대한 연구는 부재한 실정이다.

2.2 부상기술 파악 방법

기존 연구들은 부상기술의 전략적 가치를 극대화하기 위한 핵심요소로 두 가지를 제안한다. 첫째는 조기 파악, 둘째는 시

장 선점이다. 조기 파악을 위해서는 전략기획 도구, 방법론, 시스템의 고도화가 중요하다. 반면 시장 선점은 기업 내외부 다양한 이해관계자들의 상호작용과 복잡한 의사결정의 결과물이다(Haessler *et al.*, 2023). 따라서 기존 부상기술 방법론 연구들은 상대적으로 복잡도가 낮은 전자에 초점을 맞추어 왔다.

초기 부상기술 파악 방법론은 주로 전문가 평가에 의존했다. 브레인스토밍(brainstorming)과 같은 기본적인 방법들이 주로 활용되었으나, 이는 다양한 인지적 편향으로 인한 낮은 정확도가 문제였다(Hiltunen, 2008). 이에 연구자들은 신뢰성 향상을 위해 복수 전문가를 체계적으로 활용하는 방법론들을 발전시켰다. 대표적인 기법으로는 델파이 기법(Delphi method), 명목 집단화 기법(nominal group technique), AHP(analytic hierarchy process), 브레인라이팅(brainwriting) 등이 있다(Dalkey and Helmer, 1963; Delbecq and Van de Ven, 1971; Saaty, 1977). 방법별 특징은 있으나, 전문가 간 직접 상호작용을 통한 집단사고 편향의 최소화와 체계적인 의견 교환을 통해 합의를 도출한다는 것이 공통점이다. 그러나 주관적 편향과 제한된 합리성은 여전히 문제로 남아 있으며, 시간과 자원 소모가 큰 특성상 기술 수명 주기가 짧은 부상기술 파악에는 적합하지 않다(Landeta, 2006; Okoli and Pawlowski, 2004; Rowe and Wright, 1999; Tetlock, 2005; Tversky and Kahneman, 1974).

과학기술 문헌의 대규모 전자문서화, 컴퓨팅 파워의 발전, 계량서지학의 결합에 의해 탄생한 과학계량학(scientometrics)은 부상기술 파악 방법론에 새로운 전환을 가져왔다(Donthu *et al.*, 2021). 초기 연구들은 기술 성장을 유관 문서 수나 인용 수와 같은 대리 지표(proxy measure)의 변화로 측정하는 방법을 제안했다(Price, 1963). 그러나 초기 지표들은 지나치게 단순하여 측정의 정확도가 낮았다. 후속 연구는 크게 두 가지 유형으로 분류할 수 있다. 동시 인용(co-citation), 동시 출현 단어(co-word)와 같은 새로운 지표를 제시하는 연구들이 첫 번째 유형이다(Callon *et al.*, 1983; Porter and Detampel, 1995; Small, 1973). 두 번째 유형의 연구들은 인용 네트워크(citation network), 주경로(main path), 텍스트 마이닝과 같이 분석 방법 고도화에 초점을 두었다(Garfield, 1964; Hummon and Doreian, 1989; Yoon and Park, 2004). 이후 특허 지표와 머신러닝 기법을 결합하여 부상기술 식별의 정확성과 체계성을 제고하는 방향으로 방법론을 발전시켜 왔다(Lee *et al.*, 2018; Zhou *et al.*, 2021).

부상기술 파악에 있어 과학계량학 기반 방법들은 전문가 평가 기반 방법들 대비 두 가지 장점이 있다. 첫째, 대규모 데이터 분석을 통해 주관적 판단과 제한된 합리성에 의한 편향을 최소화할 수 있다. 둘째, 기술 발전의 동적 구조를 보다 명확하고 포괄적으로 파악할 수 있다. 부상기술 속성 관점에서 전자는 부상기술의 신규성과 응집성 분석의 신뢰도를 제고할 수 있으며, 후자는 성장성과 응집성의 가시화를 통해 의사결정을 용이하게 한다. 그러나 이러한 장점에도 불구하고 두 가지 근본적 한계가 있다. 서지학 기반의 어떤 지표도 부상기술

의 미래 영향력 예측에 완벽히 부합하지는 않는다(Trajtenberg, 1990). 따라서 기존 기술 발전경로를 벗어나 패러다임 전환을 가져오는 급진적 기술의 파악이 어렵다. 또한, 기술의 연구개발과 문서화 시점 간 시차(time lag)로 인해 기술 수명 주기가 짧은 부상기술을 적시에 파악하는 데는 구조적인 한계가 따른다.

이후의 연구들은 복수의 지표와 방법론을 체계적으로 조합하려는 시도가 많았다(Daim *et al.*, 2006; Érdi *et al.*, 2013; Huang *et al.*, 2014). 초기에는 과학계량학 분석을 통해 대규모 서지 데이터로부터 부상기술 후보를 추출한 후, 전문가 기반 기법들을 활용해 신규성이나 미래 임팩트를 해석, 평가하는 연구가 주류였다(Lee *et al.*, 2009; Shibata *et al.*, 2008). 그러나 이 역시 전문가 평가의 다양한 편향들이 약점으로 작용했다. 후속 연구들은 복수의 과학계량학적 방법을 체계적으로 활용해 각 방법의 단점을 보완하고 장점을 극대화하고자 하였다. 연구자들은 논문 인용 네트워크와 토픽 모델링(topic modeling), 특허 인용 네트워크와 주경로 기법, 주경로 기법과 기술 트리 등 다양한 조합의 지표와 기법들을 제시했다(Kim and Shin, 2018). 많은 시도에도 불구하고, 서지 정보만으로는 부상기술의 성장에 영향을 미치는 비언어적, 맥락적인 요소들과 불확실성을 파악, 반영하기 어렵다는 한계는 여전히 존재했다(Bradley and Lang, 1999; Esuli and Sebastiani, 2006; Hutto and Gilbert, 2014).

최근 데이터 분석 방법론 연구는 멀티모달과 인공지능 방법론에 주목하고 있다. NLP 분야에서 감성 분석과 오피니언 마이닝의 등장은 텍스트 데이터로부터 주관적 정보와 맥락적 의미를 체계적으로 포착할 수 있는 방법론을 제공하였다(Liu, 2012; Pang and Lee, 2008). 텍스트 마이닝 방법론이 내용 이해에서 감성과 맥락 분석 중심으로 발전하면서, 이미지 감성 분석도 활발하게 연구되기 시작했다(Kim, 2014; You *et al.*, 2015). 연구자들은 소셜 미디어의 이미지-텍스트 결합 데이터 분석 결과, 이미지와 텍스트의 감성이 종종 상이하다는 사실을 발견했다(Borth *et al.*, 2013; Poria *et al.*, 2015). 이를 바탕으로 텍스트, 이미지, 음성 데이터를 결합해 종합적으로 감성을 분석하는 멀티모달 분석기법이 빠르게 발전하고 있다(Baltrušaitis *et al.*, 2018; Poria *et al.*, 2017; Soleymani *et al.*, 2017; Zadeh *et al.*, 2016; Zadeh *et al.*, 2017; Zadeh *et al.*, 2018). 나아가 파운데이션 모델(foundation model)과 대조학습(contrastive learning)의 발전은 멀티모달 분석을 통한 감성 분석의 정확도를 비약적으로 향상시키고 있다(Alayrac *et al.*, 2022; Gemini Team, 2024; Jia *et al.*, 2021; Liu *et al.*, 2023; OpenAI, 2023; Radford *et al.*, 2021; Yin *et al.*, 2023).

최근 부상기술 파악 방법론 연구들은 초기의 BERT(Bidirectional Encoder Representations from Transformers)에서 최근의 ChatGPT, Gemini, LLaMA에 이르는 다양한 LLM을 기존의 지표 및 방법론과 체계적으로 조합하는 데 초점을 두고 있다(Song *et al.*, 2023; Wang *et al.*, 2024). 일부 연구가 특허에

Table 1. Emerging Technology Identification Methodologies

Classification	Methodology (selected)	Sources	Advantage	Disadvantage	References
Expert judgment	Delphi	Experts	• Tacit knowledge/ insight utilization • Groupthink avoidance (in structured methods) • Flexibility in exploring unexpected issues	• High cost • Time-consuming slow process • Various expert judgment biases • Localized knowledge (knowledge stickiness) • Limited scalability • Limited responsiveness to rapid changes	Dalkey and Helmer(1963), Delbecq and Van de Ven(1971), Linstone and Turoff(1975), Saaty(1977), Rowe and Wright (1999), Okoli and Pawlowski(2004)
	Nominal group techniques				
	AHP(analytic hierarchy process)				
	ANP(analytic network process)				
	Brainstorming				
	Brainwriting				
Scientometrics	Citation analysis	Papers, patents, proceedings, reports, news, social media contents and other scientific/technological documents	• Reduced expert judgment biases • Reduced bounded rationality • Knowledge structure identification • Dynamic knowledge evolution analysis • Excellent scalability	• Time lag (retrospective nature) • Limited future anticipation/forecasting • Limited emotional/ contextual information • Quality differences among documents • Limited responsiveness to rapid changes	Small(1973), Callon <i>et al.</i> (1983), Trajtenberg(1990), Porter and Detampel(1995), Yoon and Park(2004)
	Co-word analysis				
	Topic modeling				
	Main path				
	Social network analysis				
Systematic use of multiple methodologies	Citation network & topic modeling	Experts, documents and image	• Reduced time lag • Improved performances of future forecasting • Improved reliability of scientometrics • Consideration into quality differences among documents	• Time-consuming slow process • Limited emotional/ contextual information • Difficulty in interpreting results	Daim <i>et al.</i> (2006), Shibata <i>et al.</i> (2008), Érdi <i>et al.</i> (2013), Huang <i>et al.</i> (2014)
	Citation network & main path				
	Main path & technology tree				
	Keyword network & image similarity				
Multimodal and AI-enabled methodologies	Deep learning frameworks	Documents, image, audio and video	• Emotion/context analysis • Identification of incongruence between data of different modes	• Significant computational resources • Limited future anticipation/forecasting • Cultural biases/variances in non-verbal cues • Black-box nature of AI models	Zhou <i>et al.</i> (2021), Mariani <i>et al.</i> (2023), Gumber(2025)
	Multimodal sentiment analysis				
	Generative AI				

포함된 이미지 분석의 가치를 확인하기는 했으나, 멀티모달 분석을 통한 부상기술 파악과 속성 측정에는 이르지 못했다 (Jee *et al.*, 2022). <Table 1>에서와 같이 부상기술 파악 방법론은 전문가 판단과 과학계량학에 뿌리를 두고 있다. 두 방법론이 복수 방법론의 체계적 활용으로 수렴했고, 이제 멀티모달과 인공지능 방법론의 발전으로 새로운 전기를 맞이하고 있다. 이러한 맥락에서 본 연구는 부상기술 파악을 위해 텍스트-이미지 통합 멀티모달 분석을 시도하는 첫 연구라는 점에 의의가 있다.

3. 연구 방법

3.1 연구 프레임워크

본 연구는 영상에 포함된 전문가의 언어적 표현과 표정에서 나타나는 감성 간 불일치를 분석함으로써, 부상기술의 불확실성과 모호성을 정량적으로 측정하는 데 목적이 있다. 이에 따른 세부 분석 목표는 1) 텍스트 감성 분석, 2) 화자 표정 감성 분석, 3) 감성 분석 결과 불일치에 대한 정량적 분석으로 설정

하였다. 텍스트 감성 분석의 경우, 현재의 주류 방법론은 BERT와 LLM 계열로 양분할 수 있다. 최근 연구에 따르면, 명확한 목표하에 특정 패턴을 탐색해야 하는 경우에는 BERT 계열 방법들이 GPT-4나 LLaMA 등 대표적 LLM 대비 성능이 우수하다(Zhang *et al.*, 2025). 따라서 본 연구에서는 BERT 기반 모델을 텍스트 감성 분석에 채택했다. 화자 표정 감성 분석은 딥러닝 기반 안면 표정 인식 기법으로 감성 신호를 효과적으로 포착할 수 있음이 검증되었으나(Li and Deng, 2018), 빠르게 성능이 개선되는 분야이기 때문에 폭넓게 활용되는 방법론을 선택하기 어렵다. 따라서 연구 시점에서 강건성과 정확성이 가장 우수한 딥페이스(DeepFace) 오픈소스 플랫폼의 MTCNN(Multi-Task Cascaded Convolutional Networks)과 얼굴 표정 감성 분석 방법을 활용하였다. 이러한 딥러닝 기반 안면 인식 기술은 이미 인간 수준의 정확도에 근접했음이 입증된 바 있으며(Taigman *et al.*, 2014), 다양한 환경에서의 얼굴 인식 및 분석에 대한 높은 신뢰성을 확보하고 있다(Wang and Deng, 2021). 마지막으로 언어-표정 감성 분석 간 결과 불일치 측정 및 정량화하기 위해, 감성 불일치 가설을 이론적 기반으로 감

성 불일치 지수(Sentiment Incongruity Index, SII)를 개발하여 사용하였다.

<Figure 1>은 위의 세 방법을 중심으로 한 연구 프레임워크다. 먼저 분석 대상인 LLM의 대표적 모델들을 선정한 후, 관련 영상 데이터를 유튜브에서 수집한다. 수집된 영상에서 오디오 데이터를 추출해 텍스트 데이터로 변환하고, 토큰화 및 정규화를 포함한 기본적 전처리를 수행한다. 동시에 영상 데이터에서 얼굴 표정 영역 인식 범위와 시간 프레임을 최적화한다. 다음 단계에서는 텍스트 데이터는 BERT로, 표정 데이터는 DeepFace로 각각의 감성 분석을 수행한다. 두 감성 분석 결과 간의 불일치를 분석한 후, SII를 산출한다. 분석의 깊이를 더하고 감성 불일치 해석을 풍부하게 하기 위해 텍스트 데이터에 일반 Jaccard-cosine 유사도 2단계 군집화 기법을 적용해 LLM의 세부 핵심 기술 범주를 도출한다. 이를 바탕으로 세부 기술별 SII를 계산하고, 텍스트 및 표정 감성 분석 결과와 종합해 세부 기술 간 감성 불일치 차이를 비교 분석한다. 마지막으로 OpenAI의 ChatGPT 모델을 대상으로 기업 단위의 심층 분석을 수행한다.

Step 1. Target Technology Selection

- Define selection criteria for emerging technologies.
- Target large language models (LLMs).



Step 2. Data Acquisition and Preprocessing

- Collect expert videos from YouTube.
- Separate modalities (text vs. video frames).
- Preprocess data for each modality.



Step 3. Individual Sentiment Analysis

- Quantify modality-specific sentiment scores:
- Calculate Text Sentiment Score (TSS) via BERT.
- Calculate Facial-expression Sentiment Score (FSS) via DeepFace.



Step 4. Integrated Multimodal Sentiment Analysis

- Quantify TSS and FSS discrepancies.
- Calculate Sentiment Incongruity Index (SII).



Step 5. Clustering & In-depth Analysis

- Extract key technical topics from subtitles via TF-IDF.
- Perform 2-stage clustering to group similar topics (Jaccard-cosine).
- Analyze incongruity patterns for specific technology sub-technologies.

* BERT: bidirectional encoder representations from transformers, MTCNN: multi-Task cascaded convolutional networks, LLM: large language model, FSS: Facial-expression Sentiment Score, SII: Sentiment Incongruity Index, TF-IDF: term frequency-inverse document frequency, TSS: Text Sentiment Score

Figure 1. Overall Research Framework

3.2 감성 불일치 가설

본 연구의 방법론은 신호이론(signaling theory)에 기반한 감성 불일치 가설에 이론적 기반을 두고 있다. 신호이론은 정보 비대칭 상황에서 발신자가 신호를 보내고, 수신자가 이를 수용하는 과정을 설명한다(Spence, 1973). 주목할 점은 수신자 대비 전문성과 신뢰성이 우위인 발신자가 상대적으로 적은 비용과 노력을 들여 수신자에게 신호를 수용하게 한다는 주장이다. 신호이론에 따르면, 사회적으로 검증된 전문가의 부상기술에 대한 의견은 전문성이 비교적 낮은 사람들에게 쉽게 신뢰를 얻고 채택된다.

신호이론은 의사소통을 언어 신호가 전달되는 과정으로 간주한다. 감성 불일치 가설은 이를 확장해, 실제 의사소통은 언어뿐 아니라 다양한 비언어적 신호가 동시에 전달되는 과정이라고 주장한다. 발신자는 본인의 목적이나 수신자의 기대와 같은 다양한 요소를 고려해 언어적 신호를 의도적으로 통제한다. 그러나 비언어적 신호는 발신자가 통제하기 어렵다. 따라서 발신자는 무의식적으로 내적 감정이나 실제 의도를 비언어적 신호를 통해 표출한다(Ekman, 1972). 감성 불일치 가설은 언어적 신호와 비언어적 신호가 일치하지 않는 경우가 많다는 관찰에서 시작한다. 한 걸음 더 나아가, 언어-비언어 신호 간 불일치가 발생하는 경우 수신자들이 언어적 내용(7%)보다 어조(38%)와 표정(55%)에 비중을 두고 진의를 판단한다고 제시했다(Mehrabian and Wiener, 1967; Mehrabian and Ferris, 1967). 즉, 수신자의 수용과 해석에 있어 비언어적 신호는 간과할 수 없는 요소이다.

감성 불일치 가설 후속 연구들은 비언어적 신호의 중요성을 다양한 상황에서 관찰했다. Eisenberg(1984)는 규제·평판·법적 위험 등으로 인해 조직의 공적인 의사소통 상황에서 언어적 표현이 보수적이고 모호하게 되는 관행을 관찰했다. 반면 문화에 따른 차이는 있지만, 표정, 어조, 억양과 같은 비언어적 신호를 통해 수신자의 잠재적 기대나 발신자의 내적 동기 파악이 용이하다는 연구들이 지속적으로 보고되고 있다(DePaulo et al., 2003; Jack et al., 2012). 종합하면, 감성 불일치 가설은 발신자의 진의 파악 및 수신자의 수용성 예측을 위해 언어적 신호와 비언어적 신호를 통합적으로 파악하는 것이 필수적임을 시사한다.

3.3 데이터

본 연구는 분석 대상 부상기술로 LLM을 선정했다. GPT-4를 비롯한 LLM의 트랜스포머(transformer) 기반 모델들은 기존의 파편화된 NLP 및 비전 기술들이 하나의 거대한 파운데이션 모델로 융합되는 강한 융집성을 바탕으로 기존 모델 대비 혁신적이라고 평가받고 있으며, 모델 성능은 비약적으로 성장하고 있다(Dakhel et al., 2024; Wolf et al., 2020). 다수의 논문과 보고서들은 LLM이 2030년까지 대부분의 산업과 인간 생활에

유례없는 변화를 가져올 것이라고 전망하고 있다(McKinsey & Company, 2023; OpenAI, 2023). 그러나 이러한 급격한 성장과 광범위한 파급력 이면에는 Rotolo et al.(2015)가 부상기술의 핵심 속성으로 제시한 불확실성과 모호성이 매우 높은 수준으로 공존한다. Rotolo et al.(2015)은 불확실성을 기술 발전의 결과, 상업적 성공 가능성, 잠재적 응용 분야 등에 대한 확률적 지식이 불완전한 상태로 정의했다. 현재 LLM은 이 정의에 부합하는 세 가지 차원의 심각한 불확실성을 내포하고 있다. 첫째, 기술적 한계와 결과의 불확실성이다. 환각(hallucination)과 편향성, 블랙박스 특성으로 인한 설명 가능성 부족 등은 기술의 완성도를 저해하는 핵심 요인이다. 둘째, 상업적 성공 및 시장 생태계의 불확실성이다. 막대한 컴퓨팅 자원과 학습 비용이 요구되는 반면 명확하고 지속 가능한 수익 창출 모델은 확립되지 않았으며, 오픈소스 생태계와 빅테크 기업 간의 주도권 경쟁 결과도 예측하기 어렵다. 셋째, 잠재적 응용 분야에서의 사회·제도적 수용성 불확실성이다. 저작권 침해, 데이터 프라이버시 논란과 더불어 각국의 AI 규제 신설 등 정책이나 시장과 같은 외부환경의 변화가 기술의 확산 경로를 급격히 변경할 수 있다. 즉, LLM은 겉으로 드러나는 폭발적인 성장성에도 불구하고 기술, 시장, 제도의 다방면에서 미래 발전 경로 예측이 매우 어려운 불확실성의 집약체이다. 따라서 전문가들이 외적으로 표출하는 긍정적 전망과 내적으로 인지하는 불확실성의 괴리를 측정하는 본 연구의 방법론을 적용하기에 가장 적합한 대상이다. 또한 LLM의 미래 전망 및 주요 특성에 대한 다양한 전문가들의 영상 데이터가 풍부한 장점이 있다(Zhao et al., 2023). LLM 중에서는 기술 리더십과 시장 점유율 측면에서 명확한 우위를 가지고 있는 ChatGPT, Gemini, Claude, LLaMA를 선정했다.

LLM 관련 비디오 데이터는 유튜브에서 수집하였다. 이는 유튜브가 다양한 전문 콘텐츠 제공자들 간의 경쟁을 통해 심도 있는 기술 분석 데이터를 축적하고 있는 대표적인 시각 미디어 플랫폼이기 때문이다(Budzinski et al., 2021). LLM 4개 모델 검색 결과, 개발사 홈페이지를 제외한 대부분의 영상이 유튜브에 업로드되어 있었다. 틱톡(TikTok), 레딧(Reddit) 등 다른 플랫폼에도 일부 영상이 있었으나, 대부분 유튜브에도 업로드된 중복된 영상이었다. 수집된 데이터의 최초 업로드된 기간은 2019년 5월 1일부터 2024년 5월 30일까지이다. 해당 기간 중 위의 LLM 세부 모델 간 경쟁이 본격화되면서 기술적 혁신성, 성장성, 임팩트와 불확실성에 대한 다양한 논의가 미디어 상에서 이루어졌다. 이전 기간은 데이터가 적고, 위 기간 대비 주요 내용의 이질성이 커서 제외했다.

데이터 검색에는 LLM의 모델 키워드와 전문성 키워드를 조합하는 방식을 사용하였다. 기본 검색식은 다음과 같다: (“ChatGPT” OR “GPT-4” OR “Gemini” OR “Claude” OR “LLaMA”) (“expert” OR “technical analysis” OR “deep dive” OR “comparison” OR “critical review”). 1차로 10만 건 이상의 영상이 검색되었다. 이에 대한 전수 조사는 물리적으로 불가능하

Table 2. Overview of Collected LLM Video Data

No	LLM models	Number of videos	Average Duration (min)	Main Contents
1	OpenAI - Events and Talks	29	18	• Product/Tech Launch • Technical Demo • Keynote Speech • Case Study • Interview/Talk • Research Talk
2	Anthropic - Research	5	38	• Research Talk • Interview/Talk • Business Strategy • Education/Workshop
3	Google Gemini - Google largest and most capable AI	13	12	• Product/Tech Launch • Case Study • Technical Demo • Keynote Speech
4	Meta - Connect 2023	23	27	• Keynote Speech • Product/Tech Launch • Education/Workshop • Business Strategy • Case Study • Interview/Talk

* LLM: large language model

므로, 관련성과 파급력이 높은 영상을 선별하기 위해 2차로는 각 LLM의 공식 유튜브 계정의 LLM 관련 채널로 한정하여 개별 영상 추천 여부, 조회수, 전문적 키워드 포함 여부를 바탕으로 87편의 영상을 선정했다. 이 과정에서 전문가의 기조연설과 화자가 모두 기술 전문가로 구성된 인터뷰와 대담 위주로 선정하여, 비전문가인 진행자의 출현이 빈번하거나 다수의 패널이 동시에 등장하여 분석에 혼선을 줄 수 있는 영상은 제외했다. 또한 기술적 적합성 검증을 위해 선별된 영상에 대해 안면 인식 알고리즘(MTCNN)을 예비 적용한 결과, 분석에 오류를 유발할 수 있는 경우는 분석에서 제외했다. 구체적으로 화자의 얼굴이 직접적으로 노출되지 않는 경우와 조명 및 해상도 문제로 표정 인식이 불가능한 경우가 제외 사유에 해당된다. 마지막으로 영상 내 화자의 소속 기업 및 기관 등을 확인하여 <Table 2>와 같이 70편의 분석 대상 영상을 선정했다. 신규 모델 및 서비스를 공개하는 제품·기술 발표(Product/Tech Launch), 실제 구동 화면 시연과 기능 테스트를 하는 기술 데모(Technical Demo), 비전 및 전략 중심 발표인 기조연설(Keynote Speech), 기술 성과 및 실험 결과 공유하는 연구 발표(Research Talk), 개발자 교육, 가이드 및 튜토리얼인 교육·워크숍(Education/Workshop), 전문가 간 대화 및 질의응답인 인터뷰·대담(Interview/Talk), 수익화 및 시장 전략인 비즈니스 전략(Business Strategy), 실제 적용 사례 및 프로젝트 성과 발표인 사례 연구(Case Study) 등 주요 내용은 총 8가지로 분류할 수 있다.

분석을 위해 비디오 데이터에서 텍스트와 표정 데이터를 추

출하였다. 영문 자막이 있는 경우에는 불용어 제거와 토큰화를 포함한 기본적 전처리만 수행했다. 오디오만 있는 경우에는 자동으로 자막을 생성한 후, 오타자 수정을 포함한 전처리를 거쳐 텍스트 데이터화를 수행했다. 표정 데이터 추출은 두 단계에 걸쳐 이루어졌다.

우선 영상에서 초당 두 프레임씩을 추출해 연속된 이미지 데이터셋을 생성했다. 이는 인간의 거시 표정(macro-expression)이 통상 0.5초에서 4초간 지속된다는 Ekman(1972)의 연구 결과에 근거한 것으로, 주요 감정 변화를 놓치지 않으면서도 인접 프레임 간의 과도한 정보 중복을 제거하여 딥러닝 모델의 연산 효율성을 극대화하기 위한 조치이다. 동일한 자막이 몇 초간 지속되어 연속하는 프레임에서 얼굴 표정의 변화가 관측되는 경우도 분석에 포함되며, 호흡이나 침묵 등 자막이 없는 구간은 분석에서 제외했다.

다음에는 각 프레임에 안면 인식 알고리즘을 적용해 얼굴 영역을 확정하고, 해당 영역을 독립된 이미지로 분리해 저장했다. 저장된 표정 이미지들은 원본 영상의 타임스탬프와 동기화해 해당 시점의 텍스트에 매칭했다. 이러한 과정을 통해 각 모달리티 정보의 특성을 보존하면서, 상호 비교가 가능한 표준 데이터셋을 구축했다.

3.4 멀티모달 감성 불일치 분석

감성 불일치 분석은 세 단계로 이루어진다. 우선 BERT를 활용해 텍스트 데이터에서 텍스트 감성 점수(Text Sentiment

Score, TSS)를 도출한다. 다음으로 DeepFace를 얼굴 이미지에 적용해 표정 감성 점수(Facial-expression Sentiment Score, FSS)를 계산한다. 마지막으로 TSS와 FSS의 불일치 정도를 SII로 정량화한다.

(1) 텍스트 감성 점수

전술한 바와 같이, 본 연구에서는 텍스트 데이터 감성 분석을 위해 BERT 모델을 활용했다. 트랜스포머 기반 모델은 어텐션(attention) 메커니즘을 통해 문맥적 관계를 정교하게 파악함으로써, 기존의 순환신경망(recurrent neural network, RNN)이나 합성곱신경망(convolutional neural network, CNN) 기반 모델 대비 다양한 자연어 처리 작업에서 비약적인 성능 향상을 입증하였다(Devlin *et al.*, 2019; Vaswani *et al.*, 2017). 특히 문맥 상 단어 간 관계 파악 성능이 뛰어나기 때문에 감성 분석 정확도가 높은 장점이 있다. 최적 모델 선택을 위해 가장 널리 활용되는 플랫폼 중 하나인 허깅페이스(Hugging Face)에서 제공하는 오픈소스 라이브러리의 다양한 모델을 테스트했으며, 테스트 결과와 기존 연구 결과를 종합해 감성 분석 강건성과 정확도가 가장 높은 BERT 기반 다국어 텍스트 감성 분석 모델(bert-based-multilingual-uncased-sentiment)을 선택했다(Narayan and Kumar, 2024; NLP Town, 2025; Putra and Sibaroni, 2024; Wolf *et al.*, 2020).

$$TSS_t = \frac{s_t - 3}{2}. \quad (1)$$

본 연구에서는 초당 두 프레임으로 샘플링된 타임스탬프가 기준단위이므로, 텍스트 감성 분석도 같은 기준단위를 사용했다. 기준단위별로 1~5점 사이의 순서형 범주값 감성 점수와 신뢰도 점수를 산출했다. FSS와의 비교를 위해 식 (1)과 같이 타임스탬프 t 에서 순서형 범주값 감성 점수 $s_t \in \{1, \dots, 5\}$ 를 순서형 변환해 -1.0(매우 부정)에서 +1.0(매우 긍정)까지의 TSS를 산출했다.

(2) 표정 감성 점수

얼굴 표정 감성 분석을 위해서는 얼굴 영역 인식과 표정 감성 분석에서 현재 가용 모델 중 가장 정확도가 높은 DeepFace의 MTCNN을 채택했다. 우선 초당 두 프레임을 샘플링한 순차적 이미지 데이터에서 얼굴 영역을 인식해 이미지로 추출한다. 다음에는 얼굴 이미지를 대상으로 표정의 감성을 적용해 행복(Happy), 슬픔(Sad), 분노(Angry), 놀람(Surprise), 공포(Fear), 혐오(Disgust), 중립(Neutral)의 7가지 범주 중 하나로 분류한다. 얼굴 감성 집합 E 는 식 (2)와 같이 정의할 수 있다.

$$E = \{\text{Happy, Surprise, Neutral, Sad, Disgust, Fear, Angry}\}. \quad (2)$$

이 경우, DeepFace 분류기 C_{DF} 에 시점 t 의 얼굴 이미지 I_t 를

입력하면, 시점 t 의 표정이 특정 감성일 확률 P_t 와 예상 감성 e_t 를 다음 식 (3)과 (4)를 통해 구할 수 있다.

$$P_t(e|I_t) = C_{DF}(I_t), e \in E. \quad (3)$$

$$e_t = \underset{e \in E}{\operatorname{argmax}} P_t(e|I_t). \quad (4)$$

7개 감성 범주를 TSS와 동일한 구간에서 정량화하기 위해 Mehrabian과 Russell(1974)의 감성 모델인 PAD(Pleasure-Arousal-Dominance)를 활용했다. 기존 연구들을 참고해 본 연구는 7개 감성 범주를 PAD 모델의 -1.0(매우 불쾌)에서 +1.0(매우 유쾌) 사이의 유의가(valence) 점수로 변환했다(Bakker *et al.*, 2014). 감성 범주를 유의가로 변환하는 함수 $\phi: E \rightarrow [-1, 1]$ 의 주요 변환값은 식 (5)와 같다.

$$\begin{aligned} \phi(\text{Happy}) &= +0.9, \\ \phi(\text{Surprise}) &= +0.2, \\ \phi(\text{Neutral}) &= 0, \\ \phi(\text{Sad}) &= -0.6, \\ \phi(\text{Disgust}) &= -0.7, \\ \phi(\text{Fear}) &= -0.8, \\ \phi(\text{Angry}) &= -0.9. \end{aligned} \quad (5)$$

특정 시점(타임스탬프) t 에서 FSS는 식 (6)과 같다.

$$FSS_t = \phi(e_t). \quad (6)$$

(3) 감성 불일치 지수와 범주화

TSS와 FSS는 -1.0에서 +1.0까지의 동일 구간에서 계산된다. 따라서 불일치는 TSS와 FSS 간 거리로 계산할 수 있으며, 최소 0에서 최대 2의 값을 가진다. 예를 들어 TSS가 +1.0이고, FSS가 -1.0이면 불일치 값은 2가 된다. 그러나 최대 불일치값이 2인 경우 TSS 및 FSS의 연장선상에서 이해와 해석이 어려우므로, 0과 1 범위의 값을 가지도록 정규화해 식 (7)과 같이 특정 시점 t 에서 SII를 정의했다.

$$SII_t = \frac{|TSS_t - FSS_t|}{2}. \quad (7)$$

본 연구에서 제안한 지표들이 실제 영상 데이터에서 정량적으로 산출되는 과정을 구체화하기 위해, 수집된 데이터 중 일부 구간에 대한 분석 사례를 <Table 3>에 제시하였다. 영상의 도입부에서 타임스탬프별 자막 데이터와 이에 매칭된 화자의 표정 이미지를 동기화하여 분석한 결과, TSS와 FSS 간의 유의미한 차이가 발생하는 지점들이 식별되었으며, 이를 통해 SII가 산출되었다. 이러한 정량적 분석 결과는 화자의 언어적 표현과 비언어적 신호 사이의 유의미한 괴리가 존재함을 실증한다.

SII가 0이면 TSS와 FSS가 완전히 일치하며, 1이면 불일치가 극대화된다. 특정 시간 동안의 평균 SII는 식 (8)과 같이 계산

Table 3. Example of Multimodal Sentiment Data Extraction

Timestamp	Subtitle	Face sentiment	Text sentiment	FSS	TSS	SII
00:00:12,500	hi	neutral	very_positive	0	1	0.5
00:00:14,500	everyone hi everyone thank you thank you	sad	very_positive	-0.6	1	0.8
00:00:18,500	it's great to have you here today	sad	very_positive	-0.6	1	0.8
00:00:23,000	that's it we will start with why it's so	happy	neutral	0.9	0	0.45
00:00:23,500	that's it we will start with why it's so	happy	neutral	0.9	0	0.45
00:00:27,000	important to us to have a product that	neutral	very_positive	0	1	0.5
00:00:29,000	we can make freely available and broadly	neutral	very_positive	0	1	0.5
00:00:29,500	we can make freely available and broadly	fear	very_positive	-0.8	1	0.9

* FSS: Facial-expression Sentiment Score, SII: Sentiment Incongruity Index, TSS: Text Sentiment Score

할 수 있다. T 는 타임스탬프(t)의 합산값이며, 특정 구간의 총 영상 시간을 의미한다.

$$\overline{SII} = \frac{1}{T} \sum_{t=1}^T SII_t \quad (8)$$

주목할 SII 경계값은 0.5이다. 예를 들어 TSS가 +1.0이지만, FSS가 0이면 SII값은 0.5가 된다. 언어표현은 매우 긍정적이지만, 표정 감성은 중립적인 것이다. 즉, TSS와 FSS의 감성 범주가 변화하는 수준의 불일치가 발생한다. 따라서 SII값이 0.5 이상이면 강한 불일치(Strong Incongruence)로 범주화할 수 있다. 기존 연구들도 이와 유사한 기준으로 감성 불일치 수준을 범주화하고 있다(Ambady and Rosenthal, 1992; Mehrabian and Wiener, 1967). 특히 온라인 영상에 대한 멀티모달 분석 연구들은 감성 분석 수치 분포에 따른 범주화와 범주별 판별력에 대해 상세한 결과를 제공하고 있다(Zadeh *et al.*, 2016; Zadeh *et al.*, 2018). 이러한 기존 연구들을 참고해, 본 연구는 식 (9)와 같이 감성 불일치 수준을 SII값에 따라 거의 완전한 일치(Near-Perfect Congruence)에서 강한 불일치(Strong Incongruence)까지 6개 범주로 정의했다.

$$K(\overline{SII}) = \begin{cases} \text{Near - Perfect Congruence} & \text{if } 0.0 \leq \overline{SII} < 0.1, \\ \text{Subtle Incongruence} & \text{if } 0.1 \leq \overline{SII} < 0.2, \\ \text{Mild Incongruence} & \text{if } 0.2 \leq \overline{SII} < 0.3, \\ \text{Moderate Incongruence} & \text{if } 0.3 \leq \overline{SII} < 0.4, \\ \text{Pronounced Incongruence} & \text{if } 0.4 \leq \overline{SII} < 0.5, \\ \text{Strong Incongruence} & \text{if } 0.5 \leq \overline{SII} \leq 1.0. \end{cases} \quad (9)$$

거의 완전한 일치는 감성적 일치 상태에 해당한다. 경미한 불일치(Subtle Incongruence)와 약한 불일치(Mild Incongruence)는 TSS와 FSS의 전체 감성 범주가 같고, 약한 불일치가 있는 상태다. 보통 수준 불일치(Moderate Incongruence)와 현저한 불일치(Pronounced Incongruence)는 전체적 감성 범주는 같지만, 명확하게 불일치하는 요소나 부분들이 있는 상태이다. 전문가 영상의 특성상 Happy와 같은 극단적 긍정 감정의 출현 빈도가 낮고, 진지함을 동반한 Neutral 또는 미세한 Sad 계열의 표정이

기술적 고찰 과정에서 빈번하게 나타날 수 있다. 이는 화자의 전문성 전달을 위한 의도적 표정 제어일 가능성이 크므로, 본 연구에서는 SII가 0.5 미만인 구간을 신중한 태도로 간주하여 세분화하였고, 0.5 이상인 구간만을 실질적으로 강한 감성 불일치로 엄격하게 분류했다.

결과적으로 이렇게 구조화된 SII는 단순한 언어와 표정의 차이를 측정하는 것을 넘어, 전문가가 부상기술에 대해 인지하고 있는 불확실성과 모호성을 정량화하는 대리 지표로 기능한다. 신호이론에 기초할 때, 전문가는 사회적 기대나 긍정 편향으로 인해 언어적으로는 확신에 찬 신호(외적 기대, TSS)를 발신하기 쉽다. 하지만 해당 기술이 내포한 기술적 난제, 시장 미성숙 등 객관적 불확실성을 강하게 인지하고 있다면, 이러한 내적 불안감은 통제하기 어려운 비언어적 신호(내적 확신, FSS)로 표출될 수밖에 없다. 즉, 수리적으로 산출된 높은 SII값은 기술에 대한 외적 긍정 신호와 내적 불안 신호 간의 괴리를 뚜렷하게 포착한 결과이며, 이는 곧 해당 기술이 직면한 잠재적 위험 요인과 객관적 불확실성의 강도를 논리적으로 대변하는 근거가 된다.

(4) 감성 불일치 유형 분류

본 연구에서는 SII를 TSS와 FSS의 불일치 수준을 정량화해 화자가 이야기하는 대상에 대한 불확실성과 모호성을 측정하는 대리 지표로 제안한다. 그러나 하나의 SII값에 대해 TSS와 FSS의 긍정·중립·부정 조합은 다양하다. 따라서 정교한 감성 분석을 위해서는 TSS와 FSS를 긍정, 중립, 부정의 세 범주로 분류할 필요가 있다.

기존 연구들은 멀티모달 분석 시 감성 점수의 비교 및 범주화를 위해 식 (10)과 같은 변환을 제안하고 있다(Poria *et al.*, 2017; Zadeh *et al.*, 2016). 본 연구에서도 식 (10)의 SS 변수에 TSS 또는 FSS값을 넣어 [0,1] 구간값으로 정규화했다.

$$\widetilde{SS} = \frac{(SS - \min SS)}{(\max SS - \min SS)} \quad (10)$$

기존 연구들을 바탕으로 $\widetilde{SS} \geq 0.6$ 을 긍정(Positive, POS),

Table 4. Types of Sentiment Incongruity

No.	Text label	Face label	Incongruity type	Interpretation
1	POS	NEU	TextPos_FaceNeu	Positive text with neutral facial expression: limited optimism
2	POS	NEG	TextPos_FaceNeg	Positive text with negative facial expression: internally pessimistic
3	NEU	POS	TextNeu_FacePos	Neutral text with positive facial expression: moderately optimistic
4	NEU	NEG	TextNeu_FaceNeg	Neutral text with negative facial expression; reserved pessimism
5	NEG	POS	TextNeg_FacePos	Negative text with positive facial expression: internally optimistic
6	NEG	NEU	TextNeg_FaceNeu	Negative text with neutral facial expression: strongly limited pessimism

0.4 ≤ $\widetilde{SS}$ < 0.6 을 중립(Neutral, NEU), $\widetilde{SS}$ < 0.4 를 부정(Negative, NEG) 범주로 분류했다. 이러한 범주화는 인간의 감정 경험에 능동적 추론을 통한 범주화의 산물이라는 이론적 배경(Barrett, 2017)에 근거한다. 본 연구에서는 경계값 오분류 방식을 위해 범주별 포함 규칙을 명시적으로 적용했으며, 특정 구간에 표본이 과도하게 밀집하는 분포 왜곡이 없었기 때문에 별도의 분위수 보정을 하지 않았다. TSS와 FSS값의 감성 범주 분류결과의 가능한 조합은 <Table 4>와 같다. SII가 감성 불일치 강도를 나타낸다면, TSS-FSS 범주 조합은 감성 불일치의 방향성을 보여준다.

3.5 영상 · 텍스트 주제별 군집화

SII는 샘플링한 초당 두 프레임의 영상 단위로 계산된다. 따라서 영상 내에서 화자가 이야기하는 부상기술의 특정 주제에 대한 SII를 계산하려면, 영상을 텍스트 기반으로 주제별로 군집화해야 한다. 기존 연구들도 멀티모달 감성 분석의 유의미한 적용을 위해 주제별 또는 기술별로 데이터를 군집화했다(Gandhi *et al.*, 2023).

기존 문헌에서는 텍스트 데이터 군집화를 위해 효과성과 효율성 측면에서 다양한 방법을 제시하고 있다. 일반적으로 코사인(cosine) 유사도 기반 군집화가 군집 내 유사성을 극대화하는데 유리하며, 자카르(Jaccard) 유사도 기반 군집이 컴퓨팅 자원 소모 대비 속도가 빠르다(Manning *et al.*, 2008; Singhal, 2001). 이러한 특성을 바탕으로, 본 연구는 2단계의 순차적 군집화 방법을 적용했다. 우선 1단계에서는 일반 자카르(generalized Jaccard) 유사도 군집화를 통해 전체 데이터에서 유사 어휘들을 폭넓게 필터링해, 군집화를 위한 분석 공간을 효율적으로 축소했다. 다음 단계에서는 TF-IDF(term frequency-inverse document frequency)를 적용해 문서에서 핵심

키워드를 추출하고, 단어의 중요도를 정량화하였다(Salton and Buckley, 1988; Sparck Jones, 1972). 이후 TF-IDF의 단어 중요도를 반영한 코사인 유사도 기반 군집화를 적용해 문서 간의 유사성을 정교하게 측정, 반영해 군집을 구성했다. 이러한 2단계 접근법은 연산 자원 소요를 최소화하면서 빠른 군집화가 가능하고, 최종 군집의 응집도를 극대화할 수 있다(Li *et al.*, 2010). 마지막으로 기존 연구에서 제시한 LLM의 주요 세부 기술명을 참고해 군집명을 결정했다.

1단계 자카르 유사도 기반 군집화에서는 각 자막 텍스트를 단어 집합으로 표현하고, 두 집합 A 와 B 의 유사도를 식 (11)과 같이 단어 집합의 합집합 대비 교집합의 크기로 정의한다.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}. \quad (11)$$

선행 연구를 참고해 본 연구에서는 자카르 유사도 임계값을 0.3으로 설정했다(Zahrotun, 2016). 즉, 단어의 30% 이상을 공유하는 텍스트들을 동일 군집으로 그룹화한 것이다. 이는 군집 내 어휘 간 유사성을 확보하면서, 군집의 과도한 대형화를 방지한다.

2단계에서는 TF-IDF값을 단어의 중요도로 반영한 코사인 유사도 기반 군집화를 적용했다. 전체 자막 집합 크기 N , 자막 구간 빈도 df_i , 자막 구간 내에서의 빈도 $tf_{i,d}$ 인 경우, 자막 텍스트 구간 d 에서 단어 i 의 TF-IDF 가중치 $w_{i,d}$ 는 식 (12)와 같다.

$$w_{i,d} = tf_{i,d} \times \log \left(\frac{1+N}{1+df_i} \right). \quad (12)$$

TF-IDF 벡터 $\mathbf{x}_d = (w_{i,d}, \dots, w_{N,d})$ 들 사이의 코사인 유사도는 식 (13)과 같다.

Table 5. Core Sub-technology Clusters in LLMs

Major category	Original classification	Core technology keywords	Description
I. Foundation model	1. Architecture	model, transformer, attention, layer, embedding, positional encoding	LLM's basic structure based on Transformer architecture. The self-attention mechanism, which reflects contextual relationships, is key.
	2. Training	training, pretraining, fine-tuning, instruction tuning, dataset, RLHF, alignment	Large-scale unsupervised pretraining followed by supervised or reinforcement learning-based fine-tuning, including RLHF.

Table 5. Core Sub-technology Clusters in LLMs(Continued)

Major category	Original classification	Core technology keywords	Description
II. Core processing capabilities	3. Multimodal	vision, image, audio, video, speech, camera, sensor, voice	Extended LLMs capable of processing various inputs beyond text, such as images, audio, and video (e.g., GPT-4V, Flamingo).
	4. NLP	tokenization, token, parsing, language, translation, text generation, sentence, word	Core natural language processing technologies related to text understanding and generation. Includes word-level tokenization and various NLP tasks.
III. Application & optimization	5. UX	memory, chat, UI, app, interface, agent, demo	Conversational interfaces, long-term memory, personalized responses, and other technologies enhancing UX.
	6. Efficiency	inference, quantization, speed, latency, performance, capacity, sparsity, distillation, scalability	Efficiency-improving technologies through computational speed and parameter optimization. Includes support for low-resource devices.
IV. Responsible AI	7. Safety & alignment	bias, toxicity, alignment, misuse, trust, safety, security	Alignment and safety measures to prevent harmful outputs such as racism and sexism, ensuring responsible AI behavior.

* LLM: large language model, NLP: natural language processing, RLHF: reinforcement learning from human feedback, UI: user interface, UX: user experience

$$\cos(\mathbf{x}_1, \mathbf{x}_2) = \frac{\mathbf{x}_1 \cdot \mathbf{x}_2}{\|\mathbf{x}_1\| \|\mathbf{x}_2\|}. \quad (13)$$

코사인 유사도의 임계값은 관련 문헌을 참고하여 0.5로 설정했다. 이는 중간 수준의 유사성을 의미하는 값으로, 과도한 군집의 분할이나 대형화를 방지하는 한편, 유사한 의미의 텍스트들을 효과적으로 군집화한다(Umamaheswari and Rajesh, 2014).

마지막으로 LLM 핵심 구성요소에 관한 주요 선행 연구를 참고해, 2단계에 걸쳐 도출한 군집들의 명칭, 키워드, 의미를 <Table 5>와 같이 정리했다(Bai *et al.*, 2022; Luo *et al.*, 2024; Yin *et al.*, 2023; Zhao *et al.*, 2023).

4. 분석 결과

본 장에서는 사례 연구에 대한 멀티모달 감성 분석 결과를 제시한다. 먼저, 4개 주요 LLM 제공사의 SII 결과 분석을 통해 전반적인 경향을 파악한다. 다음으로 전체 데이터셋을 대상으로

LLM의 세부 요소 기술 클러스터별 감성 불일치 패턴을 종합적으로 분석한다. 이어서 시장을 선도하는 OpenAI 사례를 심층 분석하며, 마지막으로 도출된 감성 지표와 실제 기술 채택 성과 간의 관계를 논의한다.

4.1 LLM별 감성 불일치

4개 LLM 제공사별 평균 TSS, FSS, SII는 <Table 6>과 같다. 공통적으로 평균 TSS는 양(+)의 값을 FSS는 음(-)의 값을 가지며, SII는 Anthropic이 보통 수준인 것을 제외하고 나머지는 모두 현저한 불일치를 보이고 있다. 이는 전반적으로 전문가들이 LLM에 대해 긍정적으로 이야기하고 있지만 내적으로는 부정적인 인식을 하고 있음을 시사한다. 특히 OpenAI, Google, Meta는 SII가 상당히 높게 나타나 현저한 불일치 범주로 분류된다. 전문가들의 긍정적 발표와는 달리, 표정에서는 기술의 미래에 대한 불확실성 또는 부정적 반응이 평균적으로 강하게 나타난 것이다. 반면 상대적으로 Anthropic의 SII는 낮은 편이며 보통 수준의 불일치를 보인다. 이는 일부 명확한 불확실성 또는 위험 요인이 존재하더라도 그 강도가 타사 대비 상대적

Table 6. Multimodal Sentiment Analysis Results for LLMs (2019-2024)

LLM provider (model)	Mean TSS	Mean FSS	Mean SII (sentiment class)
OpenAI (ChatGPT)	0.312	-0.359	0.429 (Pronounced Incongruence)
Google (Gemini)	0.519	-0.055	0.409 (Pronounced Incongruence)
Anthropic (Claude)	0.227	-0.058	0.336 (Moderate Incongruence)
Meta (LLaMA)	0.439	-0.178	0.439 (Pronounced Incongruence)
Total	0.374	-0.163	0.403 (Pronounced Incongruence)

* FSS: Facial-expression Sentiment Score, LLM: large language model, SII: Sentiment Incongruity Index, TSS: Text Sentiment Score

으로 약한 것으로 간주된다고 해석할 수 있다. LLM 전반에서 공통적으로 높게 나타나는 SII는 부상기술 분석에 있어 단일 모드 분석이 갖는 한계와 멀티모달 분석의 필요성을 명확히 보여준다.

4.2 LLM 세부 기술별 감성 불일치 분석

LLM 내 세부 기술별 감성 불일치 유형과 강도를 분석하기 위해, <Table 5>의 7개 범주별 평균 SII를 산출하고 불일치 유형을 파악했다. 또한 개별 기업의 부상기술 분석 도구로서의 가치를 살펴보기 위해, LLM 기술 선도 기업인 OpenAI의 ChatGPT를 세부 기술별로 분석했다.

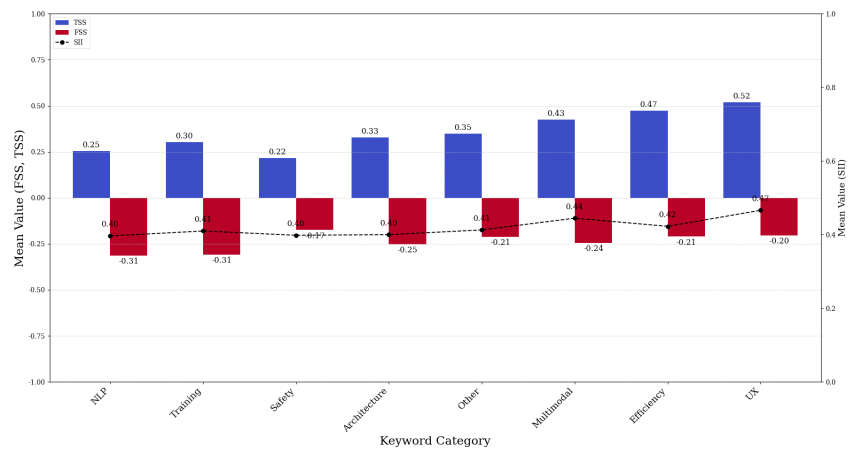
(1) 세부 기술별 감성 불일치 강도

<Figure 2>와 같이 LLM 세부 기술 전체에서 TSS는 긍정, FSS는 부정인 외적 언어표현 긍정-내적 표정 불안 조합이 일관되게

나타난다. SII는 모든 세부 기술에서 최소값 0.40 이상의 현저한 불일치를 기록했다. 특히 UX(0.47)와 Multimodal(0.44)의 SII가 가장 높았다. 즉, 전문가들은 다른 세부 기술 대비 UX와 Multimodal에서 불확실성과 위험 요인들을 인지하면서도, 대외적으로는 긍정적인 측면을 강하게 주장하고 있다. 반면 FSS 절대값이 높은 NLP(-0.31)와 Training(-0.31)에서는 불확실성과 위험 요인에 대한 인지가 UX와 Multimodal 대비 완화된 긍정적 언어 표현으로 나타난다. 따라서 TSS가 NLP(0.25), Training(0.30)으로 상대적으로 낮게 나타나는 결과로 이어진다.

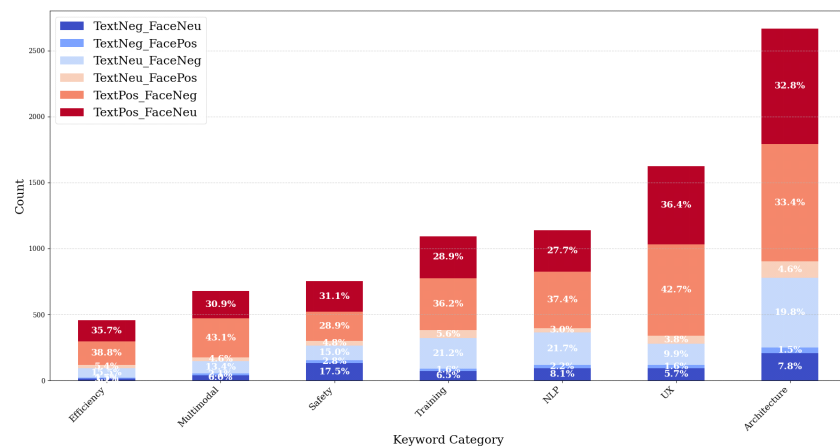
(2) 세부 기술별 감성 불일치 유형 분포

<Figure 3>에서와 같이 세부 기술 전체에서 긍정적 언어-부정적 표정(TextPos_FaceNeg)과 긍정적 언어-중립적 표정(TextPos_FaceNeu) 유형이 지배적으로 나타났다. 각 세부 기술별 두 유형의 합산 비중은 최소 60% 이상이다. 전문가들이 LLM 세부 기술을 긍정적으로 이야기하고 있으나, 내적으로는 불안감



* FSS: Facial-expression Sentiment Score, LLM: large language model, NLP: natural language processing, SII: Sentiment Incongruity Index, TSS: Text Sentiment Score, UX: user experience

Figure 2. Mean TSS, FSS, and SII by key LLM Sub-technologies



* LLM: large language model, NLP: natural language processing, UX: user experience

Figure 3. Distribution of Sentiment Incongruity Types by key LLM Sub-technologies

을 느끼고 있는 것이다. 특히 긍정적 언어-부정적 표정 조합 비중은 Multimodal(43.1%)과 UX(42.7%)에서 가장 높게 나타나, 앞선 SII 평균 분석 결과와 일치한다. 주목할 점은 세부 기술별 부정적 언어-긍정적 표정(TextNeg-FacePos) 및 부정적 언어-중립적 표정(TextNeg-FaceNeu)의 비중이다. 이 조합의 합산 비중은 Safety에서 20.3%로 가장 높게 나타난다. 즉, 전문가들이 Safety 관련 다양한 문제들의 심각성을 언어적으로는 강조하지만, 내적으로는 이를 다른 분야만큼 심각하게 받아들이지 않고 있을 가능성을 보여준다. 혹은 대중의 우려를 완화하기 위해 의도적으로 여유로운 비언어적 태도를 취하는 것으로도 해석할 수 있다.

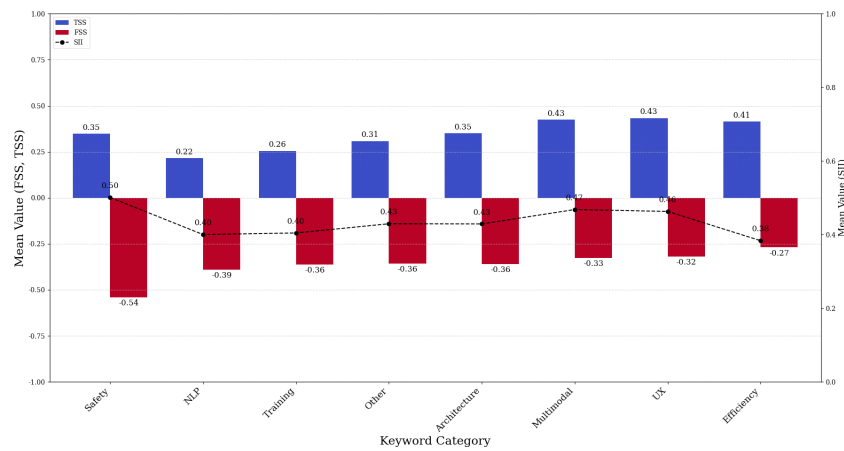
위의 결과를 종합하면, LLM의 부상기술로서의 성장에 대한 전문가들의 견해에 대해 다음과 같은 해석이 가능하다. 첫째, 전문가들은 LLM의 부상기술로서의 빠른 성장과 광범위한 영향력을 언급하고 있지만, 전반적으로 불확실성이 높다고 우려하고 있다. 둘째, 세부기술 중 Multimodal과 UX가 상대적으로 미래 불확실성이 크고 장애 요인이 될 가능성이 높다고 생각한다. 셋째,

Safety 문제에 대해서는 그 중요성이나 위험성을 언어적으로 인지하면서도, 내적으로는 이를 극복 가능한 범위 내에 있다고 보는 기술적 낙관론이 기저에 깔려 있을 가능성이 높다.

(3) OpenAI 분석

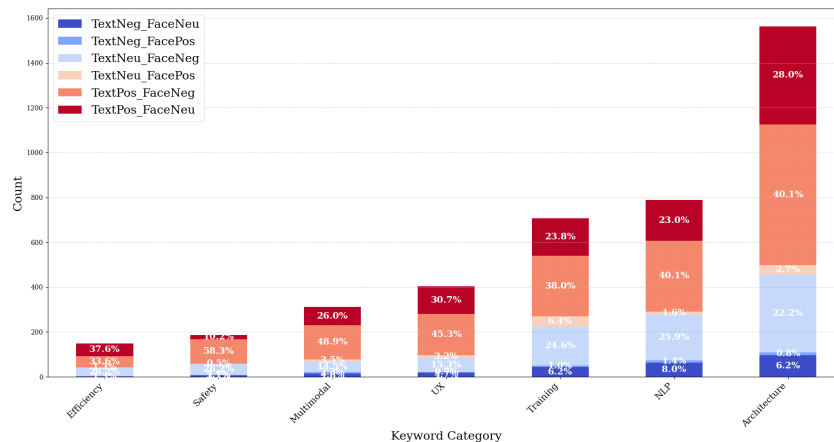
본 연구의 OpenAI 영상 분석 결과는 위의 일반적인 LLM 분석 결과와 일부 명확한 차이를 보인다. 개별 기업에 초점을 맞춰 분석할 경우, 전문가들의 특정 기업에 대한 고유한 내적 태도를 발견할 수 있는 것이다. <Figure 4>에서 OpenAI의 세부 기술별 FSS 값은 전체 평균보다 높게 나타나고 있어, 전문가들의 주요 세부 기술에 대한 내적 불안이 타사 대비 전반적으로 높다는 점을 시사한다.

특히 OpenAI Safety의 FSS(-0.54)는 4개 LLM의 Safety 기술 평균 FSS(-0.17) 대비 매우 낮으며, SII(0.50)도 평균 SII(0.40) 대비 높다. 전문가들은 타사 LLM 대비 OpenAI LLM의 Safety에 대해 상대적으로 높은 우려가 있으며, 이를 숨기기 위해 강



* FSS: Facial-expression Sentiment Score, LLM: large language model, NLP: natural language processing, SII: Sentiment Incongruity Index, TSS: Text Sentiment Score, UX: user experience

Figure 4. OpenAI: Mean TSS, FSS, and SII by key LLM Sub-technologies



* LLM: large language model, NLP: natural language processing, UX: user experience

Figure 5. OpenAI: Distribution of Sentiment Incongruity Types by key LLM Sub-technologies

한 긍정적 표현을 사용하는 경향이 있다. 실제로 OpenAI의 Safety에서 TSS(0.35)는 평균 TSS(0.22) 대비 훨씬 높게 측정되었다. 반면 SII가 평균값 대비 낮고, 내적으로 FSS가 가장 낮은 세부 기술 분야는 Efficiency이다. 전문가들의 OpenAI 최적화 및 효율화 기술에 대한 상대적으로 높은 평가를 확인할 수 있다. 요약하자면, 전문가들은 OpenAI가 평균적인 LLM들보다 불확실성이 높다고 보며, 내면적으로는 Efficiency를 강점으로, Safety를 약점으로 인식하고 있다고 할 수 있다.

<Figure 5>의 감성 불일치 유형 분석에서도 긍정적 언어-부정적 표정(TextPos_FaceNeg) 조합 비중이 Safety(58.3%)에서 가장 높게 나타나 위 분석 결과를 뒷받침한다. 이 조합 비중은 Multimodal(48.9%)와 UX(45.3%)에서 다음으로 높았다. Multimodal과 UX 영역에서 표면적으로는 긍정적 평가가 우세하게 나타났다. 그러나 표정 분석 결과에서는 전문가들의 내적 불안이 상당히 높게 나타난다. 부정적 언어-긍정적 표현(TextNeg-FacePos)과 부정적 언어-중립적 표정(TextNeg-FaceNeu) 조합의 합산 비중은 전반적으로 낮은 수준에 머물렀으나, 상대적으로 NLP(9.4%) 군집에서는 다소 두드러지는 경향을 보였다.

5. 토 의

본 연구는 영상에 포함된 전문가의 언어적 표현과 비언어적 표정 간의 감성 불일치 수준을 정량적으로 측정하는 방법과 지표를 개발하여 부상기술인 LLM에 적용하였다. 부상기술 파악과 핵심 속성 평가에 있어 가장 널리 활용되는 전문가 판단 기법들의 한계를 극복하고, 텍스트 마이닝 기법의 문제 중 하나인 긍정 편향에 대한 해결 방안을 모색하는 것이 방법론 측면의 일차적 목적이다. 특히 부상기술의 핵심 속성 중 평가가 어려운 불확실성과 모호성을 정량적으로 측정하는 데 초점을 두었다. 의사소통과 신호이론 관점에서는 언어적 신호와 비언어적 신호를 개별적으로 측정한 후 이를 비교·통합함으로써, 표면적 이성과 내적 감성의 이중적 구조를 탐색하는 새로운 방법론을 제시하고자 하였다.

공적인 의사소통 상황에서 전문가는 규제나 평판, 사회적 기대 등을 고려해 언어적 신호를 의도적으로 통제하여 긍정적인 신호를 발신할 가능성이 크다. 그러나 부상기술이 내포한 급진적 신규성이나 기술적 한계 등으로 인해 전문가가 내적으로 객관적 불확실성을 깊이 인지할 경우, 이는 인지 부조화를 유발하여 화자가 통제하기 어려운 비언어적 표정(불안, 경직 등)을 통해 무의식적으로 표출된다. 따라서 텍스트와 표정 간의 SII는 단순한 감정의 차이를 넘어, 전문가가 내적으로 감지하고 있는 인지된 불확실성과 모호성을 대변하는 정량적 대리 지표로서 강력한 논리적 타당성을 지닌다.

본 연구 결과는 언어적 신호에 있어 폴리어나 가설(Pollyanna hypothesis)에서 주장하는 긍정 편향을 뒷받침한다(Boucher

and Osgood, 1969). LLM과 세부 기술 전체에서 평균 TSS는 일관되게 양(+)의 값을 보였으며, 전체 텍스트에서 긍정적 언어 표현은 최소 60% 이상의 비중을 차지했다. 폴리어나 가설에서 주장하는 보편적 언어 표현 시의 긍정 편향이 양의 평균 TSS 값을 통해 확인된 것이다. 또한 Eisenberg(1984)의 연구와 같은 맥락에서, 공격 발표를 수행하는 전문가들이 긍정적인 표현을 중립·부정적인 표현보다 빈번하게 사용하는 경향 또한 확인되었다.

반면, 평균 FSS는 모두 음(-)의 값을 나타내며, 감성 불일치 가설에 따른 언어-비언어 신호 간의 불일치를 보여준다. 주요 4개 LLM과 세부 기술 전체에서 0.4 이상의 현저한 불일치를 보이는 SII는 이러한 경향을 보다 명확하게 나타낸다. 또한 긍정적 언어와 부정·중립적 표정의 조합 비중이 평균 60%를 상회한다는 점은, 비언어 신호에 화자의 진의가 담겨 있다는 Mehrabian과 Wiener(1967)의 주장을 뒷받침한다. 생성형 AI와 LLM은 부상기술의 성장성을 포함한 주요 속성에 대한 대규모 텍스트 분석 기반 정보를 제공하며, 다수 전문가를 대상으로 한 설문 결과 또한 방대하다. 본 연구의 결과는 이러한 정보들이 긍정적으로 편향될 확률이 높음을 시사하며, 감성적·맥락적 정보를 포함한 비언어적 신호 분석을 통한 보완의 필요성을 제기한다.

부상기술 관점에서 여섯 가지 언어-비언어 신호의 감성 조합과 SII는 새로운 분석틀을 제공한다. 우선 일치하는 기대(congruent hype) 유형이 있다. 이는 전문가들의 언어적·비언어적 표현이 일치하고 SII가 낮은 조합에 해당한다. 전문가들이 부상기술의 미래에 대해 외적으로나 내적으로 일관된 기대를 하는 경우로, 이성과 감성이 일치하는 전문가 판단으로 간주할 수 있다. 다음으로 불일치하는 기대(incongruent hype) 유형이 있다. 전문가들의 언어적 표현과 비언어적 표현이 불일치하고 SII가 높게 나타나는 사례다. 이는 전문가들의 부상기술에 대한 언어적 기대와 내적 감성이 다른 경우로, 그 판단의 신뢰성을 담보하기 어렵다. 이 중 외적 긍정-내적 불안 유형은 높은 불확실성과 위협을, 외적 부정-내적 긍정 유형은 상당한 잠재력을 내포한 것으로 해석할 수 있다.

SII는 미시적으로는 전문가 부상기술 평가의 보완지표로 활용할 수 있으며, 나아가 거시적으로는 사회-기술적 긴장(socio-technical tension) 측정 지표로도 확장 가능하다. 개별 전문가의 감성 신호 측정을 넘어, 사회 전체의 부상기술에 대한 긍정·부정의 태도나 사회-기술적 긴장 강도를 측정할 수 있는 것이다. 사회-기술적 긴장은 사회적 차원에서 기술의 혁신적 잠재력에 대한 긍정적 기대와 사회·윤리적 위험에 대한 불안이 상충할 때 발생한다(Felt and Wynne, 2007). 대규모 영상 데이터에서 긍정적 언어표현과 부정적 표정의 감성 불일치가 높은 SII로 나타난다면, 사회-기술적 긴장이 고조되고 있음을 의미한다.

본 연구는 SII를 포함한 지표와 방법론을 통해 전문가 판단의 긍정 편향과 감성 불일치를 측정할 수 있는 가능성을 입증했다. 그러나 실제 부상기술의 성장성 예측과 사회-경제적 긴

장 평가에 어느 정도 가치가 있는지는 명확하지 않다. 예컨대 SII가 가장 낮았던 Anthropic은 실제로 2025년에 안정적인 성장을 보였다. Google 및 Amazon 등으로부터 대규모 투자를 유치했으며, 시장점유율도 순탄하게 증가했다. 반면 SII가 가장 높았던 Meta의 LLaMA는 개방형 전략의 성공으로 1억 회 이상 다운로드되며 시장 점유율을 크게 높였으나, 오픈소스로 인한 보안 위험과 경쟁자 대비 약한 기술력으로 인한 기술적 성장 둔화 등의 문제가 지적되고 있다. SII가 두 번째로 높았던 OpenAI도 2023년 수입억 달러 규모의 매출을 달성하고 2025년까지 순조롭게 성장을 이어가면서, Fortune 500대 기업의 대다수가 기술을 도입하는 등 시장에서 상당한 성과를 거두었다. 그러나 Gemini 대비 경쟁 열위, 윤리 및 보안 문제, 멀티모달 데이터 한계 등이 지속적인 과제로 남아 있다.

결과적으로 SII를 포함한 본 연구의 방법론은 부상기술의 채택과 확산, 성장을 정확히 예측하는 도구로서는 아직 한계가 있다. 그러나 전문가 평가나 텍스트 기반 분석을 통한 부상기술 평가와 예측의 긍정 편향을 줄이는 데 효과적이며, 나아가 전문가의 내적 불확실성 강도 측정과 원인 탐색에도 활용도가 높다. 또한 사회-기술적 긴장을 포함한 거시지표로의 확장 잠재력도 충분하다. 부상기술 성장 다이내믹스(dynamics) 측면에서 감성 불일치가 강할수록 다양한 불확실성과 위험 요인들로 인해 기술 발전경로가 비선형적이고 역동적일 확률이 높다. 따라서 이는 기술 경로의 변동성을 예측하는 지표로서의 잠재력을 지닌다.

6. 결 론

본 연구의 학술적 기여는 크게 세 가지로 요약할 수 있다. 첫째, 전문가 평가와 텍스트 분석 중심이었던 부상기술 방법론 연구를 영상을 중심으로 한 멀티모달 분석 체계로 확장하였다. 인공지능 기반의 표정 감성 분석 방법을 도입하여 텍스트 감성 분석과 통합한 도구를 개발함으로써, 새로운 방법론 도입과 기존 방법과의 통합 측면에서 의의를 지닌다. 둘째, 전문가 평가와 텍스트 감성 분석에 대한 폴리아나 가설의 긍정 편향과 감성 불일치 가설에서 주장하는 비언어적 신호의 중요성을 정량적으로 측정하고 입증했다. 특히 비언어적 신호는 부상기술의 부정적 속성인 불확실성과 모호성에 대한 전문가의 진의를 파악하는 판단 지표로서 활용 가능하다는 측면에서 가치가 있다. 마지막으로, 최근의 대표적 부상기술인 LLM 분석을 통해, 군집화를 통한 다양한 세부 기술 단위에서의 부상기술 속성의 구체적인 분석과 기술 성장과 확산을 포함한 다이내믹스 지표로서의 확장 가능성을 제시했다.

경영 실무 측면에서는 두 가지 기여를 한다. 경영상의 부상기술 관련 의사결정에서 가장 큰 문제 중 하나는 부상기술 분석과 예측의 신뢰성이다. 본 연구에서 제안하는 지표와 방법들은 가장 널리 활용되는 전문가 평가와 텍스트 분석의 신뢰

성을 측정하고, 나아가 이를 제고할 수 있다는 점에서 일차적인 의의가 있다. 또한 부상기술에 대한 전문가의 언어적 표현과 내적 감성 분석 결과를 추적하고, 전문가에게 피드백을 제공하는 과정을 반복하면, 부상기술 평가 및 판단 역량이 높은 전문가를 육성할 수 있을 것이다.

여러 기여와 장점에도 불구하고 본 연구는 탐색적 시도라는 점에서 몇 가지 한계가 있다. 첫째, 앞서 언급한 바와 같이, 본 연구는 부상기술의 미래 경로 예측에는 활용이 제한적이다. 또한 FSS나 SII는 부상기술의 불확실성이나 모호성과 같은 핵심 속성의 대리 지표로 활용 가능하지만, 측정의 정확성을 완전히 보장하기는 어렵다. 제안한 SII는 전문가가 인지한 전반적인 불확실성의 강도를 측정하는 데 효과적이거나, 해당 불확실성이 기술적 한계에 의한 것인지, 시장 수용성 혹은 제도적 장벽에 의한 것인지에 대한 유형별 세부 분류에는 이르지 못했다. 둘째, 비언어적 신호는 다양하게 존재하나, 본 연구에서는 분석 대상을 표정 데이터로 국한했다는 점도 한계로 꼽힌다. 더욱이 전문가 영상이라는 특수한 맥락에 기인한 감성 해석은 전문가 고유의 비언어적 커뮤니케이션 스타일과 실제 기술에 대한 내적 불확실성을 완벽히 분리해 내는 데 한계가 있다. 화자가 전문성과 신뢰도를 강조하기 위해 의도적으로 취하는 진지하고 단호한 표정이나, 고도의 인지적 사고 과정에서 수반되는 무표정 등이 범용 감성 분석 모델에서는 부정적 수치로 편향되어 산출될 우려가 있다. 향후 연구에서는 이러한 상황적 진지함과 실제 내적 불확실성을 정교하게 판별하기 위해, 화자 개인의 평소 표정 패턴과 비언어적 기저 감성(baseline)을 사전에 학습한 후 특정 주제 발화 시의 상대적 편차를 분석하는 개인 맞춤형 통계 모델을 도입하여 보완해야 할 것이다. 아울러 인공지능 표정 영역 인식 및 감성 분석기법의 정확도와 같은 근본적인 기술적 문제도 존재한다. 후속 연구에서는 부상기술 분석 및 예측에서의 활용도를 높이기 위해 이러한 문제들에 대한 해결뿐 아니라, 다양한 부상기술 동적 분석기법 및 지표와의 통합도 필요하다.

참고문헌

- Alayrac, J.-B., Donahue, J., Luc, P., Miech, A., Barr, I., Hasson, Y., Lenc, K., Mensch, A., Millican, K., Reynolds, M., Ring, R., Rutherford, E., Cabi, S., Han, T., Gong, Z., Samangooei, S., Monteiro, M., Menick, J., Borgeaud, S., Brock, A., Nematzadeh, A., Sharifzadeh, S., Binkowski, M., Barreira, R., Vinyals, O., Zisserman, A., and Simonyan, K. (2022), Flamingo: A Visual Language Model for Few-Shot Learning, *Advances in Neural Information Processing Systems*, **35**, 23716-23736.
- Ambady, N. and Rosenthal, R. (1992), Thin Slices of Expressive Behavior as Predictors of Interpersonal Consequences: A Meta-Analysis, *Psychological Bulletin*, **111**(2), 256-274.
- An, H. J. and Ahn, S.-J. (2016), Emerging Technologies—Beyond the Chasm: Assessing Technological Forecasting and Its Implication for Innovation Management in Korea, *Technological Forecasting and Social Change*, **102**, 132-142.

- Bai, Y., Jones, A., Ndousse, K., Askell, A., Chen, A., DasSarma, N., Drain, D., Fort, S., Ganguli, D., Henighan, T., Joseph, N., Kadavath, S., Kernion, J., Conerly, T., El-Showk, S., Elhage, N., Hatfield-Dodds, Z., Hernandez, D., Hume, T., Johnston, S., Kravec, S., Lovitt, L., Nanda, N., Olsson, C., Amodei, D., Brown, T., Clark, J., McCandlish, S., Olah, C., Mann, B., and Kaplan, J. (2022), Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback, *arXiv preprint arXiv:2204.05862*.
- Bakker, I. C., Van Der Voordt, D. J. M., De Boon, J., and Vink, P. (2014), Pleasure, Arousal, Dominance: Mehrabian And Russell Revisited, *Current Psychology*, **33**(3), 405-421.
- Baltrušaitis, T., Ahuja, C., and Morency, L.-P. (2018), Multimodal Machine Learning: A Survey and Taxonomy, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **41**(2), 423-443.
- Barrett, L. F. (2017), The Theory of Constructed Emotion: An Active Inference Account of Interoception and Categorization, *Social Cognitive and Affective Neuroscience*, **12**(1), 1-23.
- Boon, W. P., Moors, E. H., Kuhlmann, S., and Smits, R. E. (2008), Demand Articulation in Intermediary Organisations: The Case of Orphan Drugs in the Netherlands, *Technological Forecasting and Social Change*, **75**(5), 644-671.
- Borth, D., Ji, R., Chen, T., Breuel, T., and Chang, S.-F. (2013), Large-Scale Visual Sentiment Ontology And Detectors Using Adjective Noun Pairs, *Proceedings of the ACM International Conference on Multimedia*, 223-232, ACM, Barcelona, Spain.
- Boucher, J. D. and Osgood, C. E. (1969), The Pollyanna Hypothesis, *Journal of Verbal Learning and Verbal Behavior*, **8**(1), 1-8.
- Bradley, M. M. and Lang, P. J. (1999), *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings*, University of Florida, Technical Report C-1, University of Florida, Gainesville, FL, USA.
- Budzinski, O. and Gaenssle, S. (2020), The Economics of Social Media Stars: An Empirical Investigation of Stardom, Popularity, and Success on YouTube, *Journal of Media Economics*, **32**(1-2), 1-21.
- Budzinski, O., Gaenssle, S., and Lindstädt-Dreusicke, N. (2021), The Battle of YouTube, TV and Netflix: An Empirical Analysis of Competition in Audiovisual Media Markets, *SN Business & Economics*, **1**(9), 116.
- Callon, M., Courtial, J. P., Turner, W. A., and Bauin, S. (1983), From Translations to Problematic Networks: An Introduction to Co-word Analysis, *Social Science Information*, **22**(2), 191-235.
- Christensen, C. M. (2015), *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*, Harvard Business Review Press, Boston, MA, USA.
- Cialdini, R. B. (2007), *Influence: The Psychology of Persuasion*, HarperCollins, New York, NY, USA.
- Cozzens, S., Gatchair, S., Kang, J., Kim, K. S., Lee, H. J., Ordóñez, G., and Porter, A. (2010), Emerging Technologies: Quantitative Identification and Measurement, *Technology Analysis & Strategic Management*, **22**(3), 361-376.
- Daim, T. U., Rueda, G., Martin, H., and Gerdtsri, P. (2006), Forecasting Emerging Technologies: Use of Bibliometrics and Patent Analysis, *Technological Forecasting and Social Change*, **73**(8), 981-1012.
- Dakhl, A. M., Nikanjam, A., Khomh, F., Desmarais, M. C., and Washizaki, H. (2024), *An Overview on Large Language Models, Generative AI for Effective Software Development*, 3-21, Springer Nature, Cham, Switzerland.
- Dalkey, N. C. and Helmer, O. (1963), An Experimental Application of the Delphi Method to the Use of Experts, *Management Science*, **9**(3), 458-467.
- Davis, F. D. (1989), Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, *MIS Quarterly*, **13**(3), 319-340.
- Day, G. S. and Schoemaker, P. J. (2000), Avoiding the Pitfalls of Emerging Technologies, *California Management Review*, **42**(2), 8-33.
- Delbecq, A. L. and Van de Ven, A. H. (1971), A Group Process Model for Problem Identification and Program Planning, *The Journal of Applied Behavioral Science*, **7**(4), 466-492.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., and Cooper, H. (2003), Cues To Deception, *Psychological Bulletin*, **129**(1), 74-118.
- Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. (2019), BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *Proceeding Conference North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, **1**, 4171-4186.
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., and Lim, W. M. (2021), How to Conduct a Bibliometric Analysis: An Overview and Guidelines, *Journal of Business Research*, **133**, 285-296.
- Dosi, G. (1982), Technological Paradigms and Technological Trajectories: A Suggested Interpretation of the Determinants and Directions of Technical Change, *Research Policy*, **11**(3), 147-162.
- Eisenberg, E. M. (1984), Ambiguity As Strategy In Organizational Communication, *Communication Monographs*, **51**(3), 227-242.
- Ekman, P. (1972), Universal and Cultural Differences in Facial Expression of Emotion, in J. Cole (Ed.), *Nebraska Symposium on Motivation*, **19**, 207-283, University of Nebraska Press, NE, USA.
- Érdi, P., Makovi, K., Somogyvári, Z., Strandburg, K., Tobochnik, J., Volf, P., and Zálányi, L. (2013), Prediction of Emerging Technologies Based on Analysis of the US Patent Citation Network, *Scientometrics*, **95**(1), 225-242.
- Esuli, A. and Sebastiani, F. (2006), SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining, *Proceeding International Conference on Language Resources and Evaluation*, 2200-2204.
- Felt, U. and Wynne, B. (Eds.) (2007), *Taking European Knowledge Society Seriously: Report of the Expert Group on Science and Governance*, European Commission, Brussels, Belgium.
- Festinger, L. (1957), *A Theory of Cognitive Dissonance*, Stanford University Press, Stanford, CA, USA.
- Flyvbjerg, B. (2008), Curbing Optimism Bias and Strategic Misrepresentation in Planning: Reference Class Forecasting in Practice, *European Planning Studies*, **16**(1), 3-21.
- Gandhi, A., Adhvaryu, K., Poria, S., Cambria, E., and Hussain, A. (2023), Multimodal Sentiment Analysis: A Systematic Review of History, Datasets, Multimodal Fusion Methods, Applications, Challenges and Future Directions, *Information Fusion*, **91**, 424-444.
- Garfield, E. (1964), "Science Citation Index"—A New Dimension in Indexing: This Unique Approach Underlies Versatile Bibliographic Systems for Communicating and Evaluating Information, *Science*, **144**(3619), 649-654.
- Gemini Team (Google). (2024), Gemini 1.5: Unlocking Multimodal Understanding Across Millions of Tokens of Context, *arXiv preprint arXiv:2403.05530*.
- Gumber, S. (2025), Multimodal AI Systems for Strategic Business Forecasting and Risk Management, *Journal of Data Analysis and Critical Management*, **1**(4), 35-50.
- Haessler, P., Giones, F., and Brem, A. (2023), The Who and How of Commercializing Emerging Technologies: A Technology-Focused

- Review, *Technovation*, **124**, Article 102637.
- Hiltunen, E. (2008), Good Sources of Weak Signals: A Survey of Data Sources, *Futures*, **40**(1), 21-44.
- Huang, L., Zhang, Y., Guo, Y., Zhu, D., and Porter, A. L. (2014), Four Dimensional Science and Technology Planning: A New Approach Based on Bibliometrics and Technology Roadmapping, *Technological Forecasting and Social Change*, **81**, 39-48.
- Hummon, N. P. and Dereian, P. (1989), Connectivity in a Citation Network: The Development of DNA Theory, *Social Networks*, **11**(1), 39-63.
- Hutto, C. J. and Gilbert, E. (2014), VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text, *Proceedings of the International AAAI Conference on Web and Social Media*, 215-224.
- Jack, R. E., Caldara, R., and Schyns, P. G. (2012), Culture Shapes How We Look at Faces, *Proceedings of the National Academy of Sciences*, **109**(3), 724-729.
- Jeon, J., Park, S., and Lee, S. (2022), Potential of Patent Image Data as Technology Intelligence Source, *Journal of Informetrics*, **16**(2), 101263.
- Jia, C., Yang, Y., Xia, Y., Chen, Y.-T., Parekh, Z., Pham, H., Le, Q. V., Sung, Y., Li, Z., and Duerig, T. (2021), Scaling Up Visual and Vision-Language Representation Learning with Noisy Text Supervision, *Proceedings of the International Conference on Machine Learning, PMLR*, **139**, 4916-4927.
- Kaplan, A. M. and Haenlein, M. (2010), Users of the World, Unite! The Challenges and Opportunities of Social Media, *Business Horizons*, **53**(1), 59-68.
- Kim, J. and Shin, J. (2018), Mapping Extended Technological Trajectories: Integration of Main Path, Derivative Paths, and Technology Junctures, *Scientometrics*, **116**(3), 1439-1459.
- Kim, Y. (2014), Convolutional Neural Networks for Sentence Classification, *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 1746-1751.
- Kirchberger, M. A. and Pohl, L. (2016), Technology Commercialization: A Literature Review of Success Factors and Antecedents across Different Contexts, *The Journal of Technology Transfer*, **41**(5), 1077-1112.
- Kyebambe, M. N., Cheng, G., Huang, Y. Q., He, C. H., and Zhang, Z. Y. (2017), Forecasting Emerging Technologies: A Supervised Learning Approach Through Patent Analysis, *Technological Forecasting and Social Change*, **125**, 236-244.
- Landeta, J. (2006), Current Validity Of The Delphi Method In Social Sciences, *Technological Forecasting and Social Change*, **73**(5), 467-482.
- Lee, C., Kwon, O., Kim, M., and Kwon, D. (2018), Early Identification of Emerging Technologies: A Machine Learning Approach Using Multiple Patent Indicators, *Technological Forecasting and Social Change*, **127**, 291-303.
- Lee, J., Suh, T., Roy, D., and Baucus, M. (2019), Emerging Technology and Business Model Innovation: The Case of Artificial Intelligence, *Journal of Open Innovation: Technology, Market, and Complexity*, **5**(3), 44.
- Lee, S., Yoon, B., and Park, Y. (2009), An Approach to Discovering New Technology Opportunities: Keyword-based Patent Map Approach, *Technovation*, **29**(6-7), 481-497.
- Li, S., and Deng, W. (2018), Reliable Crowdsourcing and Deep Locality-Preserving Learning for Unconstrained Facial Expression Recognition, *IEEE Transactions on Image Processing*, **28**(1), 356-370.
- Li, T., Zhang, Y., and Sindhvani, V. (2010), A Two-Stage Approach to Semi-supervised Clustering, *Proceedings of the 2010 SIAM International Conference on Data Mining*, Columbus, OH, USA, 357-368.
- Lieberman, M. B. and Montgomery, D. B. (1988), First-Mover Advantages, *Strategic Management Journal*, **9**(S1), 41-58.
- Linstone, H. A. and Turoff, M. (1975), *The Delphi Method: Techniques and Applications*, Addison-Wesley, Reading, MA, USA.
- Liu, B. (2012), *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers, San Rafael, CA, USA.
- Liu, H., Li, C., Wu, Q., and Lee, Y. J. (2023), Visual Instruction Tuning (LLaVA), *Advances in Neural Information Processing Systems*, **36**, 34892-34916.
- Luo, H., Li, P., Zhang, B., Wang, Y., Lin, C. J., and Fung, P. (2024), A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions, *arXiv preprint arXiv:2401.08984*.
- Manning, C. D., Raghavan, P., and Schütze, H. (2008), *Introduction to Information Retrieval*, Cambridge University Press, Cambridge, UK.
- Mariani, M. M., Machado, I., Magrelli, V., and Dwivedi, Y. K. (2023), Artificial Intelligence in Innovation Research: A Systematic Review, *Conceptual Framework, and Future Research Directions, Technovation*, **122**, 102623.
- Martin, B. R. (1995), Foresight in Science and Technology, *Technology Analysis & Strategic Management*, **7**(2), 139-168.
- McKinsey & Company (2023), *The Economic Potential of Generative AI: The Next Productivity Frontier*, McKinsey Global Institute, New York, NY, USA.
- Mehrabian, A. and Ferris, S. R. (1967), Inference of Attitudes from Nonverbal Communication in Two Channels, *Journal of Consulting Psychology*, **31**(3), 248-252.
- Mehrabian, A. and Russell, J. A. (1974), *An Approach to Environmental Psychology*, MIT Press, Cambridge, MA, USA.
- Mehrabian, A. and Wiener, M. (1967), Decoding of Inconsistent Communications, *Journal of Personality and Social Psychology*, **6**(1), 109-114.
- Moore, G. A. (1991), *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers*, HarperBusiness, New York, NY, USA.
- Narayan, U. and Kumar, D. (2024), Sentiment Analysis Using Transformer Based Model, *International Journal of Innovative Research in Computer Science & Technology*, **12**(1), 46-51.
- NLP Town (2025), bert-base-multilingual-uncased-sentiment [Web Page], Hugging Face, Retrieved from <https://huggingface.co/nlptown/bert-base-multilingual-uncased-sentiment>.
- Okoli, C. and Pawlowski, S. D. (2004), The Delphi Method As A Research Tool: An Example, Design Considerations And Applications, *Information & Management*, **42**(1), 15-29.
- OpenAI (2023), GPT-4 Technical Report, *arXiv preprint arXiv:2303.08774*.
- Pang, B. and Lee, L. (2008), Opinion Mining and Sentiment Analysis, *Foundations and Trends in Information Retrieval*, **2**(1-2), 1-135.
- Poria, S., Cambria, E., and Gelbukh, A. (2015), Deep Convolutional Neural Network Textual Features for Aspect-Based Sentiment Analysis, *Proceedings of the International Conference on Intelligent Text Processing and Computational Linguistics*, 42-57.
- Poria, S., Cambria, E., Hazarika, D., Mazumder, N., Zadeh, A., and Morency, L.-P. (2017), Context-Dependent Sentiment Analysis In User-Generated Videos, *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL)*, 873-883.
- Porter, A. L. and Cunningham, S. W. (2005), *Tech Mining: Exploiting New Technologies for Competitive Advantage*, John Wiley & Sons, Hoboken, NJ, USA.

- Porter, A. L. and Detampel, M. J. (1995), Technology Opportunities Analysis, *Technological Forecasting and Social Change*, **49**(3), 237-255.
- Porter, M. E. (2002), Competitiveness and Stages of Economic Development. In *Global Competitiveness Report 2001-2002 (Executive Summary)*, World Economic Forum, Geneva, Switzerland.
- Price, D. J. D. S. (1963), *Little Science, Big Science*, Columbia University Press, New York, NY, USA.
- Putra, D. F. and Sibaroni, Y. (2024), Sentiment Analysis for the 2024 Presidential Election (Pilpres) using BERT CNN, *Eduvest - Journal of Universal Studies*, **4**(11), 11042-11056.
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., and Sutskever, I. (2021), Learning Transferable Visual Models from Natural Language Supervision (CLIP), *Proceedings of the International Conference on Machine Learning*, PMLR 139, 8748-8763.
- Read, W., Robertson, N., and McQuilken, L. (2011), A Novel Romance: The Technology Acceptance Model with Emotional Attachment, *Australasian Marketing Journal*, **19**(4), 223-229.
- Rogers, E. M. (2003), *Diffusion of Innovations*, 5th ed., Free Press, New York, NY, USA.
- Rosenberg, N. (1996), Uncertainty and Technological Change, *RAND Journal of Economics*, **27**(4), 795-816.
- Rotolo, D., Hicks, D., and Martin, B. R. (2015), What Is an Emerging Technology?, *Research Policy*, **44**(10), 1827-1843.
- Rowe, G. and Wright, G. (1999), The Delphi Technique as a Forecasting Tool: Issues and Analysis, *International Journal of Forecasting*, **15**(4), 353-375.
- Saaty, T. L. (1977), A Scaling Method for Priorities in Hierarchical Structures, *Journal of Mathematical Psychology*, **15**(3), 234-281.
- Salton, G. and Buckley, C. (1988), Term-Weighting Approaches in Automatic Text Retrieval, *Information Processing & Management*, **24**(5), 513-523.
- Shibata, N., Kajikawa, Y., Takeda, Y., and Matsushima, K. (2008), Detecting Emerging Research Fronts Based on Topological Measures in Citation Networks of Scientific Publications, *Technovation*, **28**(11), 758-775.
- Singhal, A. (2001), Modern Information Retrieval: A Brief Overview, *IEEE Data Engineering Bulletin*, **24**(4), 35-43.
- Small, H. (1973), Co-citation in the Scientific Literature: A New Measure of the Relationship Between Two Documents, *Journal of the American Society for Information Science*, **24**(4), 265-269.
- Soleymani, M., Garcia, D., Jou, B., Schuller, B., Chang, S.-F., and Pantic, M. (2017), A Survey of Multimodal Sentiment Analysis, *Image and Vision Computing*, **65**, 3-14.
- Song, B., Luan, C., and Liang, D. (2023), Identification of Emerging Technology Topics (ETTs) Using BERT-based Model and Sematic Analysis: A Perspective of Multiple-field Characteristics of Patented Inventions (MFCOPs), *Scientometrics*, **128**(11), 5883-5904.
- Sparck Jones, K. (1972), A Statistical Interpretation of Term Specificity and Its Application in Retrieval, *Journal of Documentation*, **28**(1), 11-21.
- Spence, M. (1973), Job Market Signaling, *The Quarterly Journal of Economics*, **87**(3), 355-374.
- Srinivasan, R. (2008), Sources, Characteristics and Effects of Emerging Technologies: Research Opportunities in Innovation, *Industrial Marketing Management*, **37**(6), 633-640.
- Stirling, A. (2007), A General Framework for Analysing Diversity in Science, Technology and Society, *Journal of the Royal Society Interface*, **4**(15), 707-719.
- Taigman, Y., Yang, M., Ranzato, M. A., and Wolf, L. (2014), DeepFace: Closing the Gap to Human-Level Performance in Face Verification, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1701-1708.
- Teece, D. J. (2007), Explicating Dynamic Capabilities: The Nature and Microfoundations of (Sustainable) Enterprise Performance, *Strategic Management Journal*, **28**(13), 1319-1350.
- Tetlock, P. E. (2005), *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton University Press, Princeton, NJ, USA.
- Toback, E., Naudts, H., Daelemans, W., de Fortuny, E. J., and Martens, D. (2018), Belgian Economic Policy Uncertainty Index: Improvement Through Text Mining, *International Journal of Forecasting*, **34**(2), 355-365.
- Trajtenberg, M. (1990), A Penny For Your Thoughts: Patents and The Diffusion of Technological Innovations, *The Rand Journal of Economics*, **21**(1), 172-187.
- Tversky, A. and Kahneman, D. (1974), Judgment Under Uncertainty: Heuristics And Biases, *Science*, **185**(4157), 1124-1131.
- Umamaheswari, K. and Rajesh, K. (2014), An Enhanced Text Document Clustering Using Cosine Similarity and Jaccard Coefficient, *International Journal of Computer Applications*, **97**(14), 1-5.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017), Attention Is All You Need, *Advances in Neural Information Processing Systems*, **30**, 5998-6008.
- Vecchiato, R. (2012), Environmental Uncertainty, Foresight and Strategic Decision Making: An Integrated Study, *Technological Forecasting and Social Change*, **79**(3), 436-447.
- Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003), User Acceptance of Information Technology: Toward a Unified View, *MIS Quarterly*, **27**(3), 425-478.
- Wang, M. and Deng, W. (2021), Deep Face Recognition: A Survey, *Neurocomputing*, **429**, 215-244.
- Wang, Y., Liu, X., and Zhu, X. L. (2024), Enhancing Emerging Technology Discovery in Nanomedicine by Integrating Innovative Sentences using BERT and NLDA, *Journal of Data and Information Science*, **9**(4), 155-195.
- Wiefels, P. (2002), *The Chasm Companion: A Field Guide to Crossing the Chasm and Marketing High-Tech Products to Mainstream Customers*, HarperBusiness, New York, NY, USA.
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., Drame, M., Lhoest, Q., and Rush, A. M. (2020), Transformers: State-of-the-Art Natural Language Processing, *Proceeding 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 38-45.
- Yin, S., Fu, C., Zhao, S., Li, K., Sun, X., Xu, T., and Qiao, Y. (2023), A Survey of Multimodal Large Language Models, *arXiv preprint arXiv:2306.13549*.
- Yoon, B. and Park, Y. (2004), A Text-Mining-Based Patent Network: Analytical Tool for High-Technology Trend, *The Journal of High Technology Management Research*, **15**(1), 37-50.
- You, Q., Luo, J., Jin, H., and Yang, J. (2015), Robust Image Sentiment Analysis Using Progressively Trained And Domain Transferred Deep Networks, *Proceedings of the AAAI Conference on Artificial Intelligence*, 2497-2503.
- Zadeh, A., Chen, M., Poria, S., and Morency, L.-P. (2017), Tensor Fusion Network For Multimodal Sentiment Analysis, *Proceedings of the*

- Conference on Empirical Methods in Natural Language Processing*, 1103-1114.
- Zadeh, A., Zellers, R., Pincus, E., and Morency, L.-P. (2016), MOSI: Multimodal Corpus Of Sentiment Intensity And Subjectivity Analysis In Online Opinion Videos, *arXiv preprint arXiv:1606.06259*.
- Zadeh, A. B., Liang, P. P., Poria, S., Cambria, E., and Morency, L.-P. (2018), Multimodal Language Analysis In The Wild: CMU-MOSEI Dataset And Interpretable Dynamic Fusion Graph, *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 1637-1647.
- Zahra, S. A. and Nielsen, A. P. (2002), Sources of Capabilities, Integration and Technology Commercialization, *Strategic Management Journal*, **23**(5), 377-398.
- Zahrotun, L. (2016), Comparison Jaccard Similarity, Cosine Similarity and Combined Both of the Data Clustering With Shared Nearest Neighbor Method, *Computer Engineering and Applications Journal*, **5**(1), 11-18.
- Zhang, J., Huang, Y., Liu, S., Gao, Y., and Hu, X. (2025), Do BERT-Like Bidirectional Models Still Perform Better on Text Classification in the Era of LLMs?, *arXiv preprint arXiv:2505.18215*.
- Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Chen, Y., Yin, D., Li, J., Tang, Y., Li, Z., and Wen, J.-R. (2023), A Survey of Large Language Models, *arXiv preprint arXiv:2303.18223*.
- Zhou, Y., Dong, F., Liu, Y., and Ran, L. (2021), A Deep Learning Framework to Early Identify Emerging Technologies in Large-Scale Outlier Patents: An Empirical Study of CNC Machine Tool, *Scientometrics*, **126**, 969-994.

저자소개

권재현: 성균관대학교에서 기술경영전문대학원 박사과정에 재학 중이다. 연구 분야는 기술경영, 기술예측, 특허분석, 의료영상 처리이다.

신준석: 서울대학교 공과대학 산업공학과에서 기술경영 전공으로 학사, 석사, 박사학위를 받았다. 삼성경제연구소 기술산업실 연구원으로 중장기 전략 수립 및 신사업 발굴에 참여했다. 2008년부터 성균관대학교 시스템경영공학과와 기술경영전문대학원에 재직 중이다. 국내외 대기업과 기업 포사이트(corporate foresight), 기술전략, 비즈니스 모델, R&D 시스템 디자인 관련 연구, 자문 및 컨설팅을 수행했다.