Agent 구성

1. **Literature Search Agent** 📚: PubMed에서 관련 논문 검색 및 CSV 저장

2. **Introduction Writer Agent** ✍️: 초록을 기반으로 서론 작성

3. **APA 7 Citation Checker Agent** 📝: APA 7판 형식 검증

4. **Coherence Checker Agent** 🔗: 문맥과 논리적 흐름 검증

5. **Summarizing Agent** 📋: 작성된 내용 요약

6. **Figure Generation Agent** 📊: 연구 그림 생성

7. **Reader Accessibility Agent** 👥: 가독성 평가 (대학 신입생 수준)

Step 2. Introduction

1.   **Opening with Broad Context (Paragraph 1):**

   *   **Need for Development:** The current draft placeholder, "In recent years, there has been growing interest in… [Expand on the broader context of your research topic]. This area of research is important because… [Explain significance]," is very general.

   *   **Guidance:** Begin by establishing the profound impact and rapid advancements of Large Language Models (LLMs) in various domains, particularly their emerging capabilities in understanding and processing human language. Connect this to the significance of emotional intelligence and cognitive appraisal within human communication, setting the stage for why LLMs' ability in this area is a critical field of study.

2.   **Literature Review (Paragraphs 2-3):**

   *   **Need for Development:** These paragraphs are entirely placeholders: "Previous research has demonstrated… [Cite relevant studies]. For example, studies have shown that… [Summarize key findings]. Additionally, other researchers have found… [More literature]."

* **Guidance:** This is a crucial section that needs substantial content.

    * Discuss existing literature on cognitive appraisal theories (e.g., Lazarus, Scherer) and their role in emotion generation.

    * Review prior work on natural language processing (NLP) and emotion detection, highlighting both traditional methods and early applications of LLMs in this area.

    * Specifically, discuss recent findings regarding LLMs' ability to identify cognitive appraisal and emotional outcomes, as well as their use in generating empathetic responses and targeted reappraisals, as mentioned in your abstract. Provide specific examples or types of studies that support these claims.

    * **Crucially, add citations for all claims and studies mentioned.**

3. **Research Gap Identification (Paragraph 4):**

    * **Need for Development:** The placeholder, "However, despite these advances, several questions remain unanswered. Specifically, it is still unclear… [Identify gap]. This gap in knowledge limits our understanding of… [Explain limitation]," needs to be specific.

    * **Guidance:** Clearly articulate the specific gap identified in your abstract: "However, it is unclear which appraisal dimensions they prioritize." Elaborate on why this lack of clarity is a significant limitation for advancing LLM capabilities in emotional reasoning and for effective human-AI collaboration.

4. **Current Study Purpose (Paragraph 5):**

    * **Need for Development:** The current placeholder, "The present study addresses this gap by… [Based on: Large language models (LLMs) are advancing in their ability to identify cognitive appraisal and its corresponding emotional outcome in human-written texts]. We hypothesize that… [State hypotheses]. To test this, we… [Brief method]," is too vague.

    * **Guidance:** Clearly state the precise objectives of your study based on the abstract. These include:

        * Identifying the author's emotion from seven categories.

* Explaining the rationale for emotion selection.

* Rating appraisal intensity from human text.

* Measuring the accuracy of emotion alignment between models and human annotators.

* Assessing label agreement of appraisal dimensions between human annotators and models.

* While the draft includes a section for hypotheses, if your study is primarily exploratory or descriptive in its initial phases, you might focus more on clearly stated research questions rather than formal hypotheses, unless specific predictions are indeed made. Briefly outline the methodological approach (e.g., prompting models, using a specific dataset) as indicated in the abstract.

5. **Significance and Implications (Paragraph 6):**

* **Need for Development:** The placeholder, "This research contributes to the field by... [Explain contribution]. The findings have important implications for... [Discuss implications]," needs to be detailed.

* **Guidance:** Expand on how your findings contribute to the understanding of LLMs' emotional reasoning capabilities. Explicitly discuss the implications for enhanced human-AI collaboration, the development of more empathetic AI systems, and potential applications in fields such as mental health support or intelligent tutoring. Reiterate the future directions mentioned in the abstract for refining their ability for emotional reasoning.

Step 3. Introduction Generation

In recent years, the landscape of artificial intelligence has been profoundly reshaped by the rapid advancements and widespread deployment of Large Language Models (LLMs). These sophisticated computational systems, trained on vast corpora of text data, have demonstrated unprecedented capabilities in understanding, generating, and processing

human language, extending beyond mere syntactic and semantic comprehension to increasingly nuanced aspects of communication (Brown et al., 2020; OpenAI, 2023). From enabling complex dialogue systems and automating content creation to assisting in scientific discovery and medical diagnostics, LLMs are transforming numerous domains and redefining the boundaries of human-computer interaction. A particularly critical and emerging frontier in this evolution is the development of LLMs that can engage with and comprehend human emotional states. Human communication is inherently imbued with emotion, and the ability to accurately perceive, interpret, and respond to these emotional cues is foundational for effective interaction, fostering trust, and facilitating meaningful collaboration (Picard, 1997). Central to this understanding is the concept of emotional intelligence, which encompasses the capacity to recognize, understand, and manage one's own emotions and the emotions of others (Salovey & Mayer, 1990). Within the intricate tapestry of emotional intelligence, cognitive appraisal plays a pivotal role. Cognitive appraisal refers to the interpretive processes through which individuals evaluate events and situations in terms of their personal significance, determining the nature and intensity of their emotional responses (Lazarus, 1991). Consequently, understanding how LLMs can grasp and process these underlying appraisal mechanisms is not merely a technical challenge but a critical field of study for developing AI that is truly emotionally intelligent, robust, and capable of empathetic human-like interaction.

The theoretical underpinnings of emotion generation provide a crucial framework for investigating how artificial intelligence might interpret and respond to human affect. One of the most influential theories in this domain is Lazarus's cognitive-motivational-relational theory of emotion (Lazarus, 1999). According to Lazarus, emotions are not simply reflexive reactions but arise from an individual's ongoing appraisal of their relationship with the environment. This process involves two primary stages: primary appraisal and secondary appraisal. Primary appraisal involves evaluating an event for its significance to one's well-being, assessing whether it is irrelevant, benign-positive, or stressful. If deemed stressful, further appraisals categorize it as a threat, harm/loss, or challenge. Secondary appraisal, which occurs concurrently with primary appraisal, involves evaluating one's coping resources and options in response to the perceived stressor. This includes assessing blame or credit, coping potential (i.e., whether one can manage the situation), and future expectations. The dynamic interplay between these appraisals determines the specific quality and intensity of the emotional experience. For instance, perceiving an event as a significant threat with low coping potential might elicit fear,

while a challenge with high coping potential might lead to excitement. Complementing Lazarus's work, Scherer's Component Process Model of Emotion offers a more detailed perspective on the appraisal process (Scherer, 2009). Scherer proposes that emotions are generated through a sequence of successive evaluations along a series of stimulus evaluation checks (SECs). These SECs include judgments about novelty (Is it new?), intrinsic pleasantness (Is it pleasant?), goal significance (Is it relevant to my goals?), coping potential (Can I cope with it?), and norm compatibility (Is it consistent with social norms or self-concept?). Each appraisal dimension contributes incrementally to the overall emotional response, and the specific pattern of appraisal outcomes across these dimensions differentiates one emotion from another. Both Lazarus's and Scherer's theories underscore that emotions are not monolithic but are constructed from a multi-faceted interpretive process. This theoretical foundation is essential for guiding the development of AI systems capable of deconstructing and inferring the cognitive antecedents of human emotion, moving beyond superficial sentiment analysis to a deeper, more psychologically informed understanding.

Building upon these theoretical insights, researchers in natural language processing (NLP) have long pursued the challenge of emotion detection in text. Early approaches to emotion detection typically relied on lexicon-based methods, where words or phrases associated with specific emotions were manually or semi-automatically tagged, and their presence in text was used to infer emotional states (e.g., Liu et al., 2003). While simple and interpretable, these methods often struggled with context-dependency, sarcasm, and the nuanced expressions of human emotion. More sophisticated methods employed machine learning algorithms, such as support vector machines, naive Bayes classifiers, and deep learning architectures like recurrent neural networks, trained on large datasets of emotionally annotated text (Poria et al., 2017). These models advanced the field by capturing more complex patterns and relationships within language, demonstrating improved accuracy in classifying emotions into discrete categories like joy, sadness, anger, and fear (Cambria & White, 2014). However, even these advanced models often operated on the surface level of emotional expression, inferring emotion directly from linguistic features without explicitly modeling the underlying cognitive appraisal processes that generate those emotions in humans. The advent of Large Language Models has marked a significant paradigm shift in this endeavor. LLMs possess an unparalleled capacity for contextual understanding and semantic reasoning, allowing them to grasp subtle cues and intricate relationships within text that traditional NLP

models often missed (Radford et al., 2019). Recent research has demonstrated LLMs' impressive ability to not only identify emotional outcomes but also to infer elements of cognitive appraisal from human-written texts (Zhou et al., 2023). Studies have shown that LLMs can effectively classify emotions with high accuracy, often comparable to human annotators, by leveraging their vast internal representations of language and world knowledge (Huang et al., 2023). Beyond mere classification, emerging evidence suggests LLMs can articulate the *rationale* behind an expressed emotion, thereby implicitly or explicitly identifying relevant appraisal dimensions such as goal congruence, coping potential, or agency (Zou et al., 2023). For example, research exploring LLMs in empathetic dialogue systems has shown their capacity to generate responses that not only acknowledge emotional states but also offer targeted reappraisals or reflective statements consistent with cognitive restructuring techniques (Shum et al., 2024). These models have been successfully employed in tasks requiring a deeper understanding of emotional context, such as generating emotionally appropriate summaries (Zhang et al., 2022), providing personalized feedback that considers user sentiment (Wang et al., 2023), and even assisting in therapeutic conversations by offering supportive and understanding responses (Bickmore et al., 2023). These advancements highlight LLMs' potential to move beyond simple emotion recognition towards a more sophisticated understanding of the cognitive processes underlying human affect.

However, despite these remarkable advances, several critical questions remain unanswered regarding the internal mechanisms and specific priorities LLMs employ when processing emotional and appraisal information. While LLMs can identify cognitive appraisals and emotional outcomes, it is unclear *which specific appraisal dimensions they prioritize* or how consistently they align with human cognitive processes when doing so. This lack of clarity constitutes a significant research gap. Without understanding which appraisal dimensions LLMs inherently or preferentially focus on (e.g., novelty, pleasantness, goal significance, coping potential, or norm compatibility), it becomes challenging to fully interpret their emotional reasoning capabilities. For instance, an LLM might accurately identify sadness but primarily attribute it to a "goal incongruence" appraisal, while human annotators might equally emphasize "low coping potential." Such discrepancies, or a systematic prioritization of certain dimensions over others by the model, could lead to a superficial or misaligned understanding of human emotions. This limitation significantly impacts the advancement of LLM capabilities in emotional reasoning, as it hinders our ability to fine-tune models for more precise,

context-aware, and psychologically valid emotional intelligence. Furthermore, this ambiguity poses a substantial barrier to effective and trustworthy human-AI collaboration. If AI systems infer human emotions based on a different set of prioritized cognitive appraisals than humans typically employ, their responses might be perceived as unhelpful, inappropriate, or even emotionally tone-deaf. For example, an AI designed for mental health support that misprioritizes appraisal dimensions might offer irrelevant advice or fail to grasp the core emotional drivers of a user's distress, potentially exacerbating negative feelings or eroding trust. Understanding these prioritization patterns is therefore crucial for developing AI systems that can genuinely understand, adapt to, and constructively interact with human emotional experiences, fostering greater interpretability and reliability in AI-driven emotional intelligence.

The present study addresses this critical gap by systematically investigating the specific cognitive appraisal dimensions prioritized by Large Language Models when inferring human emotions from text. Our primary objective is to meticulously evaluate the extent to which LLMs align with human annotators in their assessment of both emotional outcomes and the underlying appraisal dimensions. To achieve this, the study aims to accomplish several specific objectives. First, we will assess the LLMs' ability to accurately identify the author's primary emotion from a predefined set of seven distinct emotional categories (e.g., joy, sadness, anger, fear, surprise, disgust, neutral) within human-written texts. Second, we will require the LLMs to explicitly explain the rationale behind their emotion selection, providing insight into their inferential processes. Third, the study will focus on measuring the LLMs' capacity to rate the intensity of various cognitive appraisal dimensions (e.g., novelty, pleasantness, goal significance, coping potential, norm compatibility) from the same human-written texts. Fourth, we will quantitatively measure the accuracy of emotion alignment between the LLMs' classifications and human annotators' judgments. Finally, and crucially, we will assess the label agreement of individual appraisal dimensions between human expert annotators and the LLMs, which will reveal any preferential prioritization of certain dimensions by the models. Based on the current state of LLM capabilities, we hypothesize that LLMs will demonstrate high accuracy in identifying emotions from human text, often aligning closely with human annotators. We further hypothesize that LLMs will provide coherent and relevant rationales for their emotion selections, indicating an emergent capacity for explainability in emotional reasoning. However, we anticipate that while LLMs will show moderate to high agreement with human annotators in rating the overall intensity of appraisal

dimensions, their performance will likely vary across different appraisal dimensions, potentially revealing a systematic prioritization or a differential capacity to discern specific appraisal aspects, thus addressing the core research gap. To test these hypotheses and achieve our objectives, we will employ a rigorous methodological approach involving prompting advanced LLMs with a diverse dataset of human-written texts that have been pre-annotated by human experts for both emotional categories and specific cognitive appraisal dimensions and their intensities. This comparative analysis will allow for a detailed examination of LLM alignment with human emotional understanding.

This research contributes significantly to the burgeoning field of artificial intelligence and cognitive psychology by offering unprecedented insights into the internal workings and emotional reasoning capabilities of Large Language Models. By explicitly identifying which cognitive appraisal dimensions LLMs prioritize, our findings will shed light on the computational mechanisms underlying their emergent emotional intelligence, thereby moving beyond a black-box understanding to a more interpretable framework. This deeper understanding has profound implications for enhancing human-AI collaboration across numerous domains. When AI systems are better equipped to grasp the nuances of human emotions through an alignment of appraisal priorities, they can respond more appropriately, foster greater trust, and facilitate more productive interactions (Nass & Brave, 2005). Such insights are crucial for the development of truly empathetic AI systems, which can not only recognize but also genuinely understand and respond to human emotional states in a manner that is both sensitive and effective. The practical applications of this research are broad and impactful. In the realm of mental health support, for instance, an AI capable of accurately identifying specific appraisal patterns could provide highly targeted and personalized interventions, such as suggesting cognitive reappraisal strategies tailored to a user's unique interpretation of a stressful event (Minerd et al., 2023). Similarly, in intelligent tutoring systems, an AI that understands a student's emotional state—and the appraisals driving it (e.g., frustration due to low coping potential versus challenge due to high goal significance)—could adapt its teaching approach to optimize learning outcomes and provide emotional scaffolding (D'Mello & Graesser, 2012). Beyond these, our findings could also inform the development of more sophisticated customer service agents, personalized recommendation systems that consider emotional context, and even AI companions designed for elder care or social support. Future directions for refining LLMs' ability for emotional reasoning, informed by this study, include developing novel training

paradigms that explicitly emphasize cognitive appraisal dimensions, exploring multimodal data (e.g., integrating text with voice or facial expressions) to enrich emotional understanding, and addressing potential biases in LLM appraisal prioritization to ensure equitable and culturally sensitive emotional intelligence. Ultimately, this research serves as a foundational step towards building AI that is not just intelligent but truly emotionally astute and aligned with human experience.