

사용자 문체 분석과 커스텀 GPT 구축 방안

1. 문체 분석의 핵심 요소

사용자의 글에서 **문체(writing style)**를 분석할 때는 여러 측면을 종합적으로 고려해야 합니다. 문체란 한 글쓴이가 사용하는 문장 표현상의 독특한 특징으로, 전통적으로 문장 길이, 어조, 수식의 많고 적음 등에 따라 다양한 유형으로 구분됩니다 ①. 아래에서는 문체 분석 시 핵심적으로 살펴봐야 할 요소들을 정리합니다.

1.1 기본 문체 유형

글의 전반적인 **형식과 길이**에 대한 특성입니다. 어떤 글은 **간결체**(짧고 간결한 문장 위주)인지, 혹은 **만연체**(길고 복잡한 문장 위주)인지 파악해야 합니다 ①. 예를 들어, 간결체 문장은 불필요한 단어를 생략하고 핵심만 전달하며, 만연체 문장은 수식어나 부연 설명이 풍부하여 문장이 길어집니다. 또한 글이 **서술형**(full sentences로 서술)인지, 아니면 **개조식**(핵심만 나열하거나 단문을 연속해 기술)인지도 확인합니다. 개조식 문장은 보고서 등에서 종종 사용되며 ‘~함’, ‘~음’으로 끝맺는 짧은 문장들로 구성되어 핵심을 빠르게 전달하지만, 자칫 인간미가 떨어지고 의미가 생략될 위험이 있습니다 ②. 기본 문체 유형을 파악하면 글 전반의 호흡과 템포를 이해할 수 있습니다.

1.2 말투 및 어휘 선택

말투는 글이 공식적인지 비공식적인지, 격식을 갖췄는지 혹은 친근하고 캐주얼한지를 결정하는 요소입니다. 한국어에서는 **문어체**와 **구어체**의 구분이 뚜렷합니다. 문어체는 보고서나 논문처럼 격식 있고 딱딱한 느낌으로, 어미에 “합니다”, “이다” 등을 사용하고 전문적인 한자어·격식어휘가 자주 등장합니다 ③. 반면 구어체는 일상 대화처럼 자연스럽고 편한 말투로, “해요”, “한다” 등의 표현을 쓰며 일상어와 속어도 등장하여 친근감을 줍니다 ④. 예를 들어 같은 내용을 전달해도 “본 연구에서는...”(문어체)와 “제 연구는...”(구어체)는 어조에서 큰 차이를 보입니다. 또한 **높임말(존댓말)**을 쓰는지, **반말**이나 중립어체를 쓰는지도 중요한 말투 특징입니다. 이때 대상 독자나 글의 목적에 따라 말투가 결정되며, 전문 문서일수록 경어체, 친밀한 글일수록 평어체나 구어체로 씁니다 ⑤ ⑥. **어휘 선택** 면에서는 전문용어나 한자어의 사용 빈도, 은어·속어의 사용 여부, 외국어 및 약어 사용 등도 문체의 일환입니다. 예를 들어, 기술 보고서는 전문용어와 약어를 많이 쓰고, 대중을 향한 블로그 글은 쉬운 일상어를 선호합니다. 이러한 말투와 어휘상의 특징은 글의 **톤과 독자와의 거리감**을 형성하는 핵심 요소입니다.

1.3 논리 전개 방식

글이 **어떻게 논리를 펼쳐나가는지**의 구조를 분석합니다. 크게 **연역적 전개**와 **귀납적 전개**로 나눌 수 있습니다. **연역적 전개**는 서두에 중심 주장이나 결론을 먼저 제시한 후, 이를 뒷받침하는 근거와 예시를 나열하는 방식입니다. 바쁜 독자나 전문 독자를 대상으로 한 글(예: 보고서, 기사)은 주로 두괄식(頭括式)으로 작성되어, 각 문단/문장의 처음에 요점을 배치합니다. 이에 반해 **귀납적 전개**는 여러 사실이나 사례를 먼저 제시하고 마지막에 결론을 도출하는 방식입니다. 이 방식은 독자가 글을 끝까지 읽으며 새로운 결론을 발견하도록 유도하므로, 흥미를 끌거나 스토리텔링이 필요한 글(예: 에세이, 소설)에 종종 쓰입니다 ⑦. 어떤 방식을 사용하는지는 글쓴이의 스타일과 글 목적에 따라 달라지며, **주장-근거 구조**(먼저 주장 제시 후 근거)나 **문제-해결 구조**(문제 제시 후 해결책 제안) 등도 분석 포인트입니다. 예를 들어, 블로그 글이라도 “문제가 무엇인가?”를 먼저 질문하고 뒤에 해결책을 제시하는 Q&A 형식일 수 있고, 학술적 글은 “가설-검증”의 논증 구조를 취할 수 있습니다. 또한 글이 **질문-답변(Q&A)** 형태로 전개되는지, **시간 순서대로 서술**되는지, 혹은 **대조와 비교** 구조를 취하는지도 논리 전개의 특징입니다. 이러한 전개 방식은 글이 독자에게 주장을 전달하는 **설득 전략**과도 관련이 있으므로 중요한 분석 요소입니다.

1.4 자료 인용 및 출처 표기 방식

글에서 외부 자료나 근거를 **어떻게 인용**하는지도 문체의 한 부분입니다. 예를 들어, 학술 논문이나 보고서에서는 APA나 MLA 스타일에 따라 본문 내에 괄호로 출처를 표기하거나 각주・미주를 활용합니다 8. 숫자로 된 각주[1][2]나 **괄호 내 숫자**[3][4]로 참고문헌을 표시하는 **숫자 인용 방식**도 흔합니다. 반면에 블로그나 기사에서는 **하이퍼링크**를 글 중에 삽입하거나, 문장 속에 “~라고 BBC 뉴스(2023)가 전했다”처럼 **출처명을 직접 언급**하기도 합니다. 또한 보고서나 프레젠테이션 문서에서는 도표 아래에 작은 글씨로 출처(URL 또는 기관명)를 명시하는 경우도 있습니다. 사용자가 쓴 글을 분석할 때, 출처를 **표현하는 방식**(예: “

출처

”를 붙이는지, 참고 문헌 섹션을 따로 두는지, 본문에 바로 밝히는지)을 파악하면 그 사람의 글쓰기 관습과 자료 활용 태도를 알 수 있습니다. 예컨대, 어떤 사용자는 숫자각주 대신 “(출처: ○○)” 형식으로 간략히 쓰는 버릇이 있을 수 있습니다. 인용 표기 외에도 **통계나 숫자 표현 방식**(단순 숫자 나열 vs. 표/그래프로 제시, 퍼센트 표현 등)도 자료 제시 스타일의 일부입니다. 이런 요소들을 살펴 인용과 자료 제시의 문체적 특징을 정리합니다.

1.5 문장 구조와 길이, 연결어 사용

문장 구조에서는 개별 문장이 복문인지 단문인지, 그리고 어떤 패턴을 갖추는지 살펴봅니다. 예를 들어 **단문 위주의 문체**는 하나의 문장이 하나의 아이디어만 담고 마침표로 끊어지는 형태로, 속도감 있고 명료하지만 다소 딱딱할 수 있습니다. 반면 **복문 위주의 문체**는 “~하며 ~하고 ~지만 ~다.”처럼 연결어를 많이 사용해 여러 내용을 한 문장에 담습니다. 이는 글을 풍부하고 유려하게 만들지만 잘못하면 문장이 너무 길어져 이해를 방해할 수 있습니다. 또한 문장 구조와 관련해 **수동태나 능동태** 사용 빈도도 확인합니다. 예컨대 보고서 문장에서는 “~되었다” 같은 수동 표현이 많고, 자기소개서 등에서는 “~했습니다”처럼 능동적으로 자신의 행위를 서술합니다. **연결어 사용 패턴**도 중요한데, 글에 “그러나, 그리고, 또한, 한편” 등의 접속사·전환어가 얼마나 자주 등장하는지, 문장 간 논리 관계를 어떻게 이어주는지 분석합니다. 어떤 글쓰이는 문장을 짧게 끊지만 단어만으로 이어지게(**암시적 연결**) 하고, 어떤 이는 “따라서, 즉, 게다가” 같은 명시적 연결사를 즐겨 사용합니다. 이러한 연결어 패턴은 글의 흐름과 가독성을 좌우하는 문체 요소입니다. 아울러, 문장 **길이 분포**도 살펴봐야 합니다. 예를 들어 한 문단 내 문장들이 모두 10자 내외로 짧다면 간결하고 리듬감 있는 스타일일 수 있고, 30자 이상의 문장이 많다면 설명이 자세한 만연체 경향이라 볼 수 있습니다. 요약하면, 문장 구조 분석을 통해 **문장의 호흡과 흐름**을 파악하고, 연결어와 구두점 활용을 통해 글 연결의 자연스러움과 논리성이 어떻게 표현되는지 알 수 있습니다.

1.6 어조 및 톤

어조는 글쓴이의 **태도나 감정**이 글에 배어 나오는 방식입니다. 이는 독자에게 전달되는 **분위기나 감정적 톤**과 직결됩니다. 예를 들어, **호소형 어조**는 독자에게 감정적으로 다가가 호소하거나 격려하는 톤으로, 주로 “부디 ~해 주십시오”와 같이 간곡한 부탁이나 응원의 말투가 나타납니다. **건조한 설명형 어조**는 감정을 배제하고 사실과 정보 전달에 집중하는 톤으로, 자료나 근거를 나열하며 담담하게 설명합니다. 이 경우 형용사나 감탄사의 사용이 극히 적고, 문장이 평이한 서술로 일관됩니다. **유머러스한 어조**는 독자를 즐겁게 하거나 친근감을 주기 위해 유희와 농담을 섞는 톤입니다. “이런 사소한 실수, 다들 한 번쯤은 해보셨죠?”처럼 가벼운 질문이나 말장난, 비유 등을 활용할 수 있습니다. 이 밖에도 **격양된 어조**(강조와 감탄으로 열정을 나타내는 스타일), **객관적 어조**(1인칭을 배제하고 제3자 시점으로 기술하여 객관성을 주는 스타일), **비판적 어조**(날카로운 지적과 함께 문제를 부각) 등 다양합니다. 문체 분석에서는 글 전반에 깔린 **정서적 분위기를** 파악해야 합니다. 이는 단어 선택(예: “놀랍게도”, “심각하게”), 문장 부호 사용(느낌표, 물음표의 빈도), 수사법(비유, 과장법) 등을 통해 드러납니다. 예를 들어, 느낌표와 물음표가 많은 글은 대개 격한 감정이나 반문을 담은 어조일 수 있습니다. 반면 이모티콘이나 ~~줄임표~~ 같은 표현이 많다면 캐주얼하고 농담조일 가능성이 높습니다. 글쓴이의 의도와 성격에 따라 어조는 크게 달라지므로, **설득형인지 정보 나열형인지, 친근한지 냉정한지** 등을 면밀히 살펴야 합니다.

1.7 금지 요소 및 회피하는 표현

마지막으로, 그 사람이 **일부러 피하는 문체 요소나 금기시하는 표현**이 있는지도 분석합니다. 예를 들어, 어떤 필자는 특정한 **비속어**를 절대 쓰지 않을 수 있습니다. 혹은 반대로 지나친 **미사여구**(지나치게 화려한 수식)를 경계하여 사용하지 않는 습관이 있을 수 있습니다. 비즈니스 문서나 공식 글에서는 일반적으로 “~너~”, “~할게요” 같은 지나치게 구어적인 표현이나 사투리를 피하며, 이모티콘, 과도한 느낌표, 음슴체 등을 금기시합니다. 또한 글쓰기가 1인칭이나 2인칭을 의도적으로 회피하는지도 살펴봅니다. 예컨대 학술 문체는 “~라고 할 수 있다”처럼 **자의적인 판단 표현**이나 “나는 ~라고 생각한다” 같은 **주관적 진술**을 피하고, 보다 중립적인 표현을 쓰는 경향이 있습니다. 정책 보고서 등에서는 “반드시 ~해야만 한다”와 같은 단정적인 어투를 삼가고 “~가 바람직할 것으로 판단된다”처럼 완곡하게 표현하기도 합니다. **금지 요소**에는 이 밖에도 금기어(특정 집단 비하 표현, 차별적 언어 등)의 배제, 지나친 감정 표현 억제 등이 포함됩니다. 요컨대, 글쓰기가 어떤 표현을 **일부러 쓰지 않으려 애쓰는지**를 보면 그 사람의 문체 기준과 가치관을 엿볼 수 있습니다. 이러한 요소를 정리하면, 커스텀 GPT가 글 생성 시 피해야 할 표현이나 반드시 지켜야 할 어조의 가이드라인으로 활용할 수 있습니다.

2. 글 목적 및 대상에 따른 문체 조정 요소

문체는 글이 **어떤 목적으로** 쓰였고 **누구를 대상으로** 하는지에 따라 적절히 조정되어야 합니다. 동일한 사람이라도 SNS에 올리는 글과 공식 보고서에서의 문체는 크게 다릅니다. 이 섹션에서는 다양한 글의 목적/장르별로 요구되는 문체상의 조정 요소와 특징을 살펴보고, 한 글쓰기가 **기본 문체의 일관성**을 유지하면서도 상황에 맞게 어조와 표현을 변화시키는 기준을 설명합니다.

2.1 SNS 게시글 문체

SNS(소셜 미디어) 글은 보통 **불특정 다수**의 대중에게 공개되며, 즉각적인 **호응과 공유**를 목적으로 합니다. 따라서 문체가 **친근하고 경쾌**하게 구성되는 경향이 있습니다. 문장은 인쇄 매체보다 **훨씬 짧고 간결**하게 쓰이는 것이 좋습니다 ⁹. 이는 모바일 화면 등에서 빠르게 읽혀야 하기 때문입니다. 실제로 웹상의 글은 종이 문서보다 읽는 속도가 느리므로, 보통 인쇄물의 절반 길이로 내용을 압축하라고 권장됩니다 ⁹. 예를 들어 블로그 포스트나 페이스북 글에서는 긴 문단보다는 한두 문장 단위의 짧은 문단, 혹은 리스트 형식으로 가독성을 높입니다. 말투는 **문어체보다 구어체**에 가깝게, 마치 독자와 대화하듯이 씁니다. “~입니다” 대신 “~이에요”나 “~네요” 등 부드러운 어미를 쓰고, “여러분도 느껴보셨나요?”처럼 독자에게 말을 거는 형식도 흔합니다. 다만 가벼움과 전문성의 균형을 고려해 **지나친 속어 남발은 피**하고 적절한 **높임말**을 섞어 쓰기도 합니다 ¹⁰. 어휘는 일상적인 표현을 쓰되, 해시태그나 이모티콘, 느낌표!!! 등을 활용하여 감정을 표현하고 **시각적 임팩트**를 줍니다. 예를 들어 “최고입니다”처럼 감정을 아이콘으로 나타내거나, “대박! 정말 놀랍네요!”처럼 말맛을 살립니다.

구조적으로는, **중요한 내용을 앞부분에 배치**하는 것이 효과적입니다. SNS 글은 매우 많은 콘텐츠 속에서 첫 몇 초 만에 승부가 갈리기 때문에, 첫 문장이나 제목에 독자의 시선을 끌 키워드를 넣고 핵심 메시지를 먼저 제시합니다 ¹¹. 이를 위해 헤드라인을 달거나 글의 **첫 문장을 강렬하게** 작성하는 전략이 필요합니다. 이후 본문에서는 중요한 정보를 우선 서술하고 상세 내용은 뒤에 두는 **도치된 구조**(inverted pyramid 구조)를 사용하면 독자 유지에 유리합니다 ¹¹. 형식 면에서 **표, 이미지, 동영상 등 멀티미디어**를 적극 활용하는 것도 SNS 문체의 특징입니다. 단조롭게 글만 나열하기보다 관련 사진이나 짤, 간단한 표 emoji 등을 넣어 **시각적 흥미**를 유발합니다 ¹². 예를 들어 트위터에서는 짧은 문장 + 이미지 1장을 첨부해 메시지를 극대화하고, 인스타그램에서는 해시태그와 이모티콘으로 감성을 표현합니다. 이러한 멀티모달 요소는 문체의 일부로 간주될 수 있습니다.

마지막으로, SNS 글은 **참여와 대화**를 목표로 하므로 **행동 유도 문구(CTA)**나 질문을 종종 넣습니다. “댓글로 의견을 남겨주세요!” 같은 문장은 독자의 반응을 자극하여 쌍방향 소통을 유도합니다. 요약하면 SNS 문체는 **짧고 쉬운 문장, 구어체 말투, 감정 표현이 풍부한 어조**를 특징으로 하며, **키워드 중심**으로 정보를 배치하고 **시각 요소**를 가미하는 방향으로 조정됩니다 ¹³ ¹⁴.

2.2 정보전달형 블로그 문체

블로그 글은 SNS보다 비교적 긴 호흡을 가지며, **특정 주제에 대한 정보 전달**을 주요 목적으로 합니다. 독자는 관심 있는 정보를 얻기 위해 찾아오는 경우가 많으므로, 문체는 **친절하고 체계적인** 것이 바람직합니다. 말투는 **너무 딱딱하지 않은 공식적 어조**를 사용하되, **전문성**을 살짝 가미하는 정도로 조절합니다. 예컨대 개인 블로그라도 기술/IT 주제의 포스팅이라면 “~입니다” 체로 정중하게 쓰면서도, 적절히 1인칭을 써서 경험담을 전하기도 합니다. 즉, **일인칭 시점과 독자에 대한 이인칭 호칭**(예: 여러분) 등을 활용하여 대화하듯 진행하면서, 전체적인 톤은 신뢰감을 주도록 지나치게 가법지 않게 유지합니다. 정보전달형 블로그의 문체 핵심은 **명료함과 친근함의 균형**입니다.

구조적으로는 **두괄식 전개**를 많이 택합니다. 글 도입부에 “이번 글에서는 ~~에 대해 알아보겠습니다”처럼 **주제를 선언**하고, 본문에서는 소제목(H2, H3 등)을 나누어 체계적으로 정보를 전합니다. 문단도 짧게 나누고, **번호나 Bullet 리스트**를 활용하여 핵심 내용을 정리합니다. 이러한 **시각적 구분**은 온라인 독자들의 읽기 습관에 맞춰 가독성을 높여 줍니다¹². 또한 검색엔진 최적화(SEO)를 고려해 중요한 키워드(예: “문체 분석”, “GPT 커스텀”)를 적절히 반복 배치하는 것도 블로그 글 문체의 숨은 요소입니다. 다만 이는 글의 내용 범위를 벗어나지 않는 선에서 자연스럽게 이뤄져야 하므로, 문체 분석 시에는 해당 글이 **키워드 위주로 간결하게 쓰였는지** 또는 **장황하게 서술되었는지** 살펴볼 수 있습니다.

어휘 선택 면에서는, 대상 독자층에 따라 **전문용어 사용 수준**을 조절합니다. 정보전달형 글이라 해도, 독자가 입문자이면 쉬운 용어와 풀이를 곁들이고, 전문 독자면 관련 업계 용어를 그대로 사용해 깊이를 더합니다. 이러한 톤 조절 예로, 같은 의학정보를 전달해도 일반인을 대상으로 하면 “고혈압(혈압이 높은 상태)”처럼 설명을 붙이고, 전문가 대상이면 “수축기/이완기 혈압 수치”처럼 전문용어 중심으로 서술합니다.

또한 블로그 글은 **신뢰성 확보**도 중요하여, 적절한 **출처 표시와 인용**이 문체에 녹아 있습니다. 예를 들어, “한 연구에 따르면(서울대 OO연구, 2022)...”처럼 본문 중에 자료 출처를 간략히 밝히거나, 글 말미에 참고 링크 목록을 추가하는 식입니다. 이러한 출처 제시 방식을 문체 분석에서 파악하면 블로그 글의 신뢰도와 스타일을 알 수 있습니다.

전반적으로 정보전달형 블로그 문체는 **독자를 배려한 쉽게 읽히는 구조, 필요한 만큼의 격식, 그리고 체계적인 내용 구성**이 특징입니다. 이는 SNS 글보다 차분하지만 학술 글보다는 자유로운 중간 지대에 해당하며, 커스텀 GPT는 이러한 미묘한 균형을 잘 잡도록 학습되어야 합니다.

2.3 대중 연설문 문체

대중 연설문(예: 축사, 강연, 정치 연설)은 글이 쓰여진 후 **말로 전달되는 것**을 전제로 하기 때문에, 문체가 **구어적이면서도 격식**을 갖춘 독특한 형태입니다. 연설문은 **청중의 관심을 즉각적으로 사로잡고 감동을 주는 것**이 중요하므로, 문체의 톤은 **설득력 있고 활기차며 명료**해야 합니다. 보통 **문장 길이는** 너무 길지 않게 끊어서, **호흡 단위로 전달**되도록 합니다. 글을 읽는 대신 듣는 청중을 고려해, 하나의 문장에 한 메시지만 담고, 문장 사이에 적당한 **숨과 강세**를 둘 수 있게 구성합니다. 예를 들어 연설문 초반부는 “안녕하십니까, 여러분.”처럼 짧은 인사말과 함께 시작하고, 핵심 주장을 소개하는 **도입부 한 문장**으로 청중의 이목을 집중시킵니다¹⁵.

연설문의 문체에는 **청중을 직접 지칭**하고 호흡을 같이 하는 표현이 자주 등장합니다. “친애하는 시민 여러분,” “여러분도 아시다시피” 등으로 청중에게 말을 거는 형태로 친밀감을 형성합니다. 또한 **수사학적 기교**가 많이 활용되는데, 그 중 하나가 **반복과 열거**입니다. 예를 들어 “우리는 싸웠습니다. 우리는 이겨냈습니다. 그리고 우리는 앞으로 나아갈 것입니다.”처럼 유사한 문장 구조를 반복해 리듬감을 주고 감정을 고조시킵니다. **3의 법칙**도 연설문에서 흔한데, 중요한 메시지를 세 부분으로 나누어 제시함으로써 기억에 남게 하는 방식입니다(예: “자유를 위해, 평화를 위해, 미래를 위해”). 실제로 효과적인 연설문의 표준 구조는 서론에서 청중 관심을 끌 한 문장, 본문에서 세 가지 핵심 사항, 결론에서 요약으로 이루어진다고 합니다¹⁵.

어조 면에서는 연설자는 **열정적이면서 진정성 있는 톤**을 유지해야 합니다. 문체 분석 시 연설문 특유의 **격양된 어조나 호소하는 어투**가 나타나는지 확인합니다. 예를 들어 감탄문이나 수사 의문문(“과연 이것이 정의입니까?”)을 통해 청중의 감정을 흔드는 경우가 많습니다. 또한 **직설적**으로 “~해야 합니다”처럼 강하게 촉구하거나, 반대로 **포용적**으로 “함께

나아갑시다”처럼 청중을 포함한 어구를 사용하기도 합니다. 이러한 어조 선택은 연설의 목적(설득, 계몽, 단결 등)에 따라 달라지지만, 공통적으로 **듣는 이의 심리를 고려한 리듬과 감정**이 중요합니다.

연설문 문체의 또 다른 특징은 **이야기나 비유 사용**입니다. 추상적 개념을 전하기보다, 짧은 일화나 비유를 통해 청중의 이해와 공감을 얻습니다. 문체 분석 시 연설문에 이러한 **스토리텔링 요소**가 있는지 보면 글쓰기가 대중에게 다가가는 방식을 알 수 있습니다.

마지막으로, 연설문은 **문장 부호**나 서식에서도 특징이 있는데, 실제 발화 시의 멈춤이나 강조를 표시하기 위해 줄바꿈이나 점을 이용합니다. 예컨대 원고 상에서 한 문장을 여러 줄에 걸쳐 끊어 적어 호흡을 조절하고, 중요한 단어는 대문자나 강조 표기로 표시하는 경우도 있습니다 (예: “**함께**”). 이처럼 대중 연설문 문체는 말하기 특성을 고려한 **구어체+문어체 혼합**이라 볼 수 있으며, 커스텀 GPT는 이러한 요소(청중 지칭, 반복, 3부 구조, 호소력 등)를 반영하여 연설체 문장 생성이 가능해야 합니다 16 15 .

2.4 공식 보고서 문체

공식 보고서나 공문서는 **형식성과 정확성**이 최우선인 글로, 문체가 극도로 **격식 있고 객관적**입니다. 이러한 문서에서는 개인적 감정이나 불필요한 수식이 배제되고, 사실 전달과 논리 전개가 명확하게 드러나야 합니다. **경어체**는 조직 문화에 따라 다르지만 일반적으로 “~합니다” 체 또는 **평서형** “~다” 체로 쓰며, 1인칭 “나”나 2인칭 “당신”은 등장하지 않습니다. 대신 **3인칭**이나 **수동형** 표현이 많아, 문장이 다소 딱딱하고 추상적으로 느껴질 수 있습니다 17 . 예를 들어, “우리는 ~하였다”보다는 “~한 것으로 분석되었다”처럼 **행위의 주체를 숨기는** 표현이 흔합니다. 이는 보고서가 개인이 아닌 조직의 공식 입장을 담는 문서이기 때문입니다.

문장 구조 상 특징적인 것은, **개조식 문장**의 활용입니다. 개조식은 문장의 일부 요소(주어 또는 불필요한 조사 등)를 생략하고 **요점만 간추려 항목별로 나열**하는 형식을 말합니다 18 . 예컨대 일반 서술형으로 “올해 상반기 매출은 10% 증가하였으며 비용은 5% 감소하였다.”라고 쓸 것을, 개조식으로는 “- 상반기 매출: 전년 대비 10% 증가\n- 상반기 비용: 전년 대비 5% 감소” 처럼 표기합니다. 보고서에서는 이런 개조식 나열을 통해 핵심 정보를 한눈에 들어오게 합니다. 또한 문장의 종결도 “~이다”나 “~한다” 대신 **명사형(~함, ~음)**으로 끝맺는 전형적인 보고체를 쓰기도 합니다 2 . 예를 들어 결론 부분에 “종합하면 다음과 같은 시사점을 얻을 수 있음.”처럼 끝맺습니다. 이러한 표현은 역사적으로 일제시대 권위적 문체의 잔재라는 지적도 있지만 19 , 오늘날까지 공공기관 문서에 많이 남아 있습니다.

| 장점 | 단점 |
|---|---|
| <ul style="list-style-type: none"> • 내용을 쉽게 이해할 수 있다. • 요점을 빨리 찾을 수 있다. • 필요한 항목만 볼 수 있다. • 요점을 추가하거나 삭제하기 쉽다. • 빠진 항목을 확인하기 쉽다. | <ul style="list-style-type: none"> • 생략하거나 축약하면 의미를 왜곡시킬 여지가 있다. • 인간미가 없다. • 쌍방향 소통에 부적합하다. |

개조식 문체의 장단점을 보여주는 표입니다. 왼쪽 ‘장점’에는 “내용을 쉽게 이해할 수 있다”, “요점을 빨리 찾을 수 있다”, “필요한 항목만 볼 수 있다” 등 **개조식의 효율성**이 정리되어 있습니다. 오른쪽 ‘단점’에는 “생략하거나 축약하면 의미를 왜곡시킬 여지가 있다”, “인간미가 없다”, “쌍방향 소통에 부적합하다” 등 **개조식의 한계**가 적혀 있습니다. 실제로 개조식은 정보를 빠르게 전달하기엔 좋지만 서술형 문장보다 딱딱하고 정서적 호소력이 떨어지는 경향이 있습니다. 따라서 보고서 문체에서는 개조식과 서술형을 적절히 혼용하여, **핵심은 개조식으로 간략히, 보충 설명은 서술형으로** 풀어 쓰는 것이 권장됩니다 20 21 .

보고서 문체의 또 다른 특징은 **두괄식**과 **객관식** 서술입니다. 문단의 첫 문장에 결론이나 요점을 배치하고 이하 세부사항을 나열하는 방식으로 작성하여, 바쁜 의사결정자가 **빠르게 핵심 파악**을 할 수 있도록 합니다. 또한 개인 의견을 배제하기 위해 “~라고 생각한다” 대신 “~로 판단된다”와 같이 **완곡한 단정**을 사용합니다. 수치와 근거를 많이 포함하되, 그 출

처를 각주나 본문에 정확히 명시하여 신뢰성을 부여합니다. 예를 들어 “시장 점유율은 45%로 나타났다(자료: ○○기관 보고서)”처럼 쓰는 식입니다.

전반적으로 공식 보고서 문체는 **일정한 형식(template)**과 톤을 유지하는 것이 중요합니다. 일관된 용어 사용(같은 사안을 지칭할 때 단어를 통일), 문법 오류 제로, 어휘의 정확한 의미 전달 등이 엄격하게 지켜져야 합니다^{17 22}. 문체 분석 시, 해당 글이 이런 규범을 잘 따르고 있는지 (예: 동일 용어 반복, 능동태/수동태의 통일성, 축약어 정의 등) 확인합니다. 또한 **표나 그림의 사용 여부**도 살펴봅니다. 보고서는 텍스트만 빼곡하게 쓰기보다는 중요한 데이터는 표/그래프로 제시하고 본문에서 “<표 1> 참조” 식으로 언급합니다. 글쓴이가 보고서 문체에 익숙하다면 이러한 시각 자료 활용 및 지시 표현이 나타날 것입니다.

마지막으로, **결론 도출 방식**도 보고서 문체의 포인트입니다. 결론 부분에서 “따라서 ~을 제안한다.”, “요약하면 ~이다.”처럼 명료하게 마무리 짓는지, 아니면 여운을 남기는지를 봅니다. 공식 문서라면 보통 할 말을 다 하고 난 뒤 “이상입니다.”로 끝내는 등 깔끔히 맺습니다. 이렇듯 공식 보고서 문체는 **간결함, 객관성, 형식 통일**이 특징이며, 커스텀 GPT는 이러한 요소들을 반영하여 **한글 보고서** 문장을 생성할 수 있어야 합니다.

2.5 자기소개서/이력서 문체

자기소개서(자소서)나 이력서는 **자신을 어필**하는 글이지만 동시에 격식을 갖춰야 하는 글입니다. 문체의 **목적**은 지원자의 개성과 역량을 전달하면서도 **신뢰감과 성실함**을 느끼게 하는 것이므로, **솔직하고 긍정적인 톤**이 요구됩니다. 일반적으로 1인칭 “저는”을 사용하여 자신의 경험을 서술하되, 지나치게 가볍거나 자기 자랑만 늘어놓는 어투는 피합니다. **어조**는 공손하고 정중해야 하므로 “~했습니다”, “~합니다”와 같은 **존댓말**로 일관합니다. 반말이나 지나친 감탄, 유머는 금물입니다. 대신 **단정적인 표현**과 **긍정 어휘**를 활용합니다. 예를 들어 “게으른 편입니다”보다는 “효율적인 시간관리를 위해 꾸준히 노력해왔습니다”처럼 긍정적으로 포장하는 식입니다.

문체 면에서 **간결하고 논리 정연함**이 특히 중요합니다. 인사담당자는 많은 지원자의 글을 짧은 시간에 검토하기 때문에, **길지 않은 문장으로 핵심만 담는 문체**가 좋습니다²². 실제로 자소서 작성 요령에서는 “과다한 수식어나 추상적 표현을 피하고, 짧은 단문 위주로 쓰라”고 조언합니다²². 문장이 길어지면 산만해지고 읽는 사람이 부담을 느끼므로 한 문장을 40~50자 이내로 끊는 것이 바람직합니다. 또한 **문장을 능동태로** 쓰고 **시제를 일관되게** 유지하는 것도 중요합니다. 가령 과거 경험은 “~했습니다”로 서술하고, 현재의 강점은 “~합니다”로 기술하며, 혼용하지 않도록 합니다²³. 능동태와 피동태를 혼동해서 쓰면 문장의 힘이 떨어지므로, “~을 달성했습니다”처럼 자신의 역할을 분명히 드러내는 표현이 좋습니다²³.

개성과 진솔함도 자기소개서 문체의 한 요소입니다. 수백 통의 천편일률적 자소서 속에서 눈에 띄려면 약간의 **독창적 구성**이나 **흥미 요소**가 필요합니다²⁴. 물론 형식의 틀을 깨진 않아야 하지만, 예컨대 글의 시작을 “어릴 적 제게는 두 가지 별명이 있었습니다. 바로 ‘호기심 천국’과 ‘해결사’였습니다.”처럼 흥미로운 일화로 열면 읽는 이를 끌어들이는 효과가 있습니다. 문체 분석 시 지원자가 어떤 **후크(hook)**를 사용했는지, 개성을 드러내는 **키워드**(예: 좌우명, 가치관 등)를 어떻게 배치했는지 등을 볼 수 있습니다²⁵. 다만 개성을 추구하더라도, **지나친 유행어**나 가벼운 어투는 삼가야 합니다. “~해서 개꿀이었습니다” 같은 표현은 물론 NG이며, “~한 것 같습니다... ~했던 것 같아요”처럼 자신 없는 어투도 좋지 않습니다. **자신감 있지만 겸손한** 표현을 찾는 것이 핵심입니다.

금지 요소로는 **과장과 거짓**, 그리고 **부정적 서술**이 있습니다. 문체 자체로는 이것이 드러나지 않을 수 있으나, 예를 들어 지나치게 화려한 미사여구(“탁월한 리더십으로 모두를 완벽히 이끌었습니다”)는 오히려 진정성을 떨어뜨립니다²⁶. 문체 분석 단계에서, 지원자가 얼마나 **솔직담백한 어조**를 썼는지 확인해야 합니다. “할 수 없습니다” 같은 부정문보다는 “~을 배우게 되었습니다”처럼 긍정문 위주로 서술하는 것이 일반적입니다²⁶. 또한 자신의 약점을 쓰더라도 문장을 부정적으로 끝내지 않고 “~을 개선하기 위해 ~~하고 있습니다”처럼 **발전형 어조**로 마무리합니다. 이러한 부분이 글에 나타나면 문체상의 **긍정 프레임** 유지 전략으로 볼 수 있습니다.

마지막으로, 자기소개서는 지원하는 기업이나 기관에 맞춰 **맞춤형**으로 문체를 조정하기도 합니다. 지원 기업이 보수적인 조직이면 더 격식을 차린 문체를, 창의성을 중시하는 업종이면 약간 자유로운 표현을 쓸 수도 있습니다²⁵. 이는 “일관된 문체 유지”와 상충되지만, 큰 줄기는 유지한 채 단어 선택이나 강조점만 바꾸는 식으로 해결합니다. 예를 들어 기본

문체가 담담하고 논리적인 사람이라면, 광고회사 지원용 자소서에서는 담담함 속에 창의적 사례 하나를 어조만 살짝 바꿔 소개하는 식입니다. 결국 자소서 문체는 **간결·명확, 긍정·자신감, 진솔·개성**이라는 세 축을 잘 살리는 방향으로 조정되며, 커스텀 GPT도 이러한 가이드라인에 따라 글을 생성하도록 학습시킬 수 있습니다 22 26 .

2.6 문체 일관성 유지와 상황별 변용

한 글쓴이가 여러 종류의 글을 작성할 때, **기본 문체의 일관성**을 유지하면서도 글 목적에 맞게 어조와 표현을 변형하는 것은 중요합니다. 예를 들어 평소 글이 유머러스하고 비유를 잘 쓰는 사람이 있다고 합시다. 이 사람이 **SNS 글**을 쓸 때는 그 장점을 살려 재치있는 밈(meme)이나 짧은 농담을 덧붙일 것입니다. 그러나 같은 사람이 **공식 보고서**를 작성할 때는, 유머러스한 비유 대신에 좀 더 절제된 언어로 핵심만 전달할 것입니다. **핵심 스타일**인 “어려운 개념을 쉬운 비유로 설명하는 능력”은 유지하되, 보고서에서는 농담조 대신 알기 쉬운 현실 사례나 비유 하나 정도를 간략히 언급하는 선으로 조절하는 식입니다. 즉, **내용 전달 방식**의 개성은 살리면서 **형식과 어조**는 해당 장르의 규범에 맞게 바꾸는 것입니다.

일관된 문체를 유지한다는 것은, 독자가 “이 글은 ~~답다”라고 느낄 수 있는 **특유의 품격**이나 **흐름**이 유지된다는 뜻입니다. 예를 들어 어떤 저자는 문서 종류와 상관없이 **논리적이고 차분한 전개**를 선호한다면, SNS 글에서도 감정을 과하게 넣기보다는 차분한 어휘를 쓰고, 연설문을 작성하더라도 흥분을 자제하며 논리적 설득에 집중할 것입니다. 반대로 창의적이고 감성적인 문체를 가진 사람은 보고서를 쓰더라도 다소 문학적인 표현이나 스토리텔링을 가미할 수 있습니다 (물론 과하지 않은 범위 내에서). 이처럼 **문체의 기본 성향**(논리 vs 감성, 직설 vs 완곡 등)은 그대로 유지하되, **겉으로 드러나는 형태**(말투, 형식)는 상황에 따라 변형됩니다.

변화의 기준은 **독자와 목적**입니다. 글쓴이는 각 상황의 독자가 다르다는 것을 염두에 두고, 독자가 기대하는 **톤앤매너**에 맞춰줍니다. 이를테면, 유튜브 대본을 쓸 때는 영상 시청자에게 말을 거는 2인칭과 가벼운 농담을 써도 되지만, 동일 주제로 신문 기고문을 쓸 때는 3인칭 객관 서술과 격식을 차리는 식입니다. **같은 사람이 쓴 글이라도**, “누구에게 말하느냐”에 따라 표현만 달리하는 것이죠. 그러나 **문장 구성 습관**이나 **어휘 선택의 버릇**은 완전히 숨기기 어렵습니다. 이를테면 긴 문장을 좋아하는 사람은 SNS에서도 조금은 길게 쓰는 경향이 있고, 짧은 문장을 즐기는 사람은 보고서에서도 군더더기 없이 쓰는 경향이 있습니다. 이러한 습관이 바로 그 사람의 **문체적 지문(指紋)**이고, 일관성의 핵심입니다.

예를 들어, 한 글쓴이가 일상 블로그에서는 “~했어요, 정말 재미있었어요!” 같은 **캐주얼한 어조**로 쓰지만, 그 사람의 글에는 항상 **비교를 통한 설명**이 들어있다고 가정합시다. 이 사람이 공식 자료를 작성할 때는 “~에 비유하자면 ~와 같습니다.”처럼 격식을 차려 같은 비교를 할 것입니다. 겉보기엔 하나는 반말+감탄, 다른 하나는 경어+진지함으로 다르지만, **비교를 활용해 설명하는 스타일** 자체는 변함이 없습니다. 이처럼 스타일의 기저 요소는 유지하면서, 표현 수위만 조절하는 것이 문체 변용의 기본 원칙입니다.

문체 분석을 통해 이러한 **일관성과 변용 패턴**을 파악하면, 커스텀 GPT를 설계할 때 큰 도움이 됩니다. GPT 모델에게 사용자 문체의 **불변 요소**(예: “항상 단문 위주로 씀”, “비유를 잘 활용함”)와 **가변 요소**(예: “격식 정도는 글의 포맷함에 따라 조정”)를 가르쳐줄 수 있기 때문입니다. 실제 적용에서는, 시스템 프롬프트나 fine-tuning 데이터에 “이 사용자는 어떤 글을 쓰든 ~한 경향을 유지하지만, 공식적인 글일수록 ~한 어조를 쓴다”와 같은 규칙을 넣어 모델이 상황에 맞게 스타일을 조정하도록 할 수 있습니다. 결국 목표는 **동일 필체**를 유지하면서도 **컨텍스트 적응**이 가능한 AI 글쓰기이며, 이를 위해 문체의 일관된 특징과 변용 기준 모드를 데이터화해야 합니다.

3. GPT 적용을 위한 문체 구조화 방식

사용자 문체를 분석해낸 다양한 요소들을 **GPT 모델에 적용**하려면, 우선 이러한 정보를 명시적으로 **구조화**해서 모델이 이해하도록 해야 합니다. 즉, 사용자의 글에서 추출한 문체 특징들을 일종의 **프로필 데이터**로 정리하고, 이를 GPT에게 알려주거나 학습시키는 과정이 필요합니다. 이 절에서는 문체 특징을 태깅/구조화하는 방법, JSON·YAML 등의 포맷 예시, 그리고 실제 사용자 글을 분석하여 학습용 입력 데이터로 변환하는 예를 살펴보겠습니다.

3.1 문체 요소 태깅 및 분류

먼저 앞서 정의한 문체 요소들을 체계적으로 태그합니다. 각 요소별로 미리 정해진 카테고리나 스펙트럼이 있을 수 있습니다. 예를 들어:

- **base_style**: 글의 기본 문체 유형 - 간결체 / 만연체 / 서술형 / 개조식 등.
- **formality**: 말투의 공식적 정도 - 격식/공식적 / 반격식/일상적.
- **perspective**: 시점과 높임말 - 1인칭 존댓말 / 3인칭 객관.
- **vocabulary_level**: 어휘 수준 - 전문용어 많음 / 일반어 중심 / 속어 포함.
- **tone**: 어조/톤 - 건조한설명 / 열정적호소 / 유머러스 / 격양됨 등.
- **logic_style**: 전개 방식 - 두괄식 / 미괄식 / 문답법 / 비유와스토리.
- **sentence_length**: 문장 길이 - 단문 위주 / 복문 많음.
- **connectives_usage**: 연결어 사용 - 잦음 / 최소한.
- **citation_style**: 인용 표기 - 숫자각주 / 본문출처표기 / 하이퍼링크.
- **forbidden_elements**: 피하는 요소 - 예: 구어적 표현 없음, 비속어 없음, '~한다' 체 사용 등.

이런 식으로 **키-값 쌍**으로 문체 특징을 표현할 수 있습니다. 필요에 따라 세분화된 카테고리를 더 둘 수도 있습니다 (예: tone을 세부적으로 friendly, sarcastic, formal, etc.로 분류).

문체 분석을 통해 얻은 결과를 이러한 태그로 **정형화**하면, GPT 모델이 이를 입력으로 받아들이고 글 생성 시 참조할 수 있습니다. 특히 **fine-tuning(파인튜닝)**을 한다면, 입력 데이터로 “[스타일 프로필] + [콘텐츠]”를 주고 출력으로 “[해당 스타일로 쓴 글]”을 내도록 훈련시킬 수 있습니다. 또는 **프롬프트 엔지니어링**으로, 사용자의 문체 태그를 시스템 메시지에 삽입해 둔 뒤 모델이 답변을 생성하게 할 수도 있습니다.

3.2 JSON/YAML 구조화 예시

위에서 논의한 태그들을 실제 활용할 때는 JSON, YAML 등 기계가 읽기 좋은 포맷으로 구조화합니다. 아래는 예시입니다 (JSON 형식):

```
{
  "author_id": "user123",
  "style_profile": {
    "base_style": "간결체",
    "formality": "격식",
    "perspective": "1인칭 존댓말",
    "vocabulary_level": "일반어 중심",
    "tone": ["논리적", "차분함"],
    "logic_style": ["두괄식", "주장-근거"],
    "sentence_length": "단문 위주",
    "connectives_usage": "최소",
    "citation_style": "본문출처표기",
    "forbidden": ["속어", "이모티콘", "추상적 미사여구"]
  }
}
```

위 JSON은 가상의 **author_id=user123**에 대한 문체 프로필을 나타낸 것입니다. **style_profile** 내에 해당 사용자의 문체 분석 결과를 키-값으로 모두 넣었습니다. 예를 들어 **base_style**이 "간결체"로 되어 있어 이 사용자는 군더더기 없이 간략히 쓰는 편이며, **tone**에 ["논리적", "차분함"]이 있어 감정보다는 논리 전달에 중점을 둔 어조임을 알 수 있습니다. **forbidden** 목록에는 이 사용자가 피하는 요소로 속어, 이모티콘 등이 열거되어 있습니다.

또한 YAML 형식으로 표현하면 가독성을 높일 수 있습니다. 같은 정보를 YAML로 표현한 예시는 아래와 같습니다:

```
author_id: user123
style_profile:
  base_style: "간결체"
  formality: "격식 (합니다체)"
  perspective: "1인칭 (존댓말)"
  vocabulary_level: "일반어 위주"
  tone:
    - "논리적"
    - "차분함"
  logic_style:
    - "두괄식"
    - "주장-근거 전개"
  sentence_length: "단문 위주"
  connectives_usage: "연결어 최소 사용"
  citation_style: "본문 내 출처 표기"
  forbidden:
    - "속어 사용"
    - "이모티콘"
    - "추상적 미사여구"
```

YAML은 들여쓰기를 통해 구조를 표시하므로 사람도 한눈에 보기 쉽습니다. JSON/YAML 중 어떤 포맷을 사용할지는 시스템 구현이나 선호에 따라 결정됩니다. 중요한 것은 **문체 특성이 빠짐없이 표현**되고, 나중에 GPT에게 해당 정보를 명시적으로 넘겨줄 수 있다는 점입니다.

이러한 구조화 예시는 GPT 모델의 입력 컨텍스트에 포함되어야 합니다. 예컨대 OpenAI API를 사용할 경우, 프롬프트에 `<style_profile>...</style_profile>` 같은 형태로 XML/JSON을 삽입해도 되고, 시스템 메시지에 "여기 user123의 문체 특성이 있다. 이걸 따라 작성해라"고 자연어로 써줄 수도 있습니다. 하지만 자연어 지시는 모델이 해석을 잘못할 가능성이 있으므로, 위와 같이 **포맷화된 프로필**을 두고 이를 파싱하는 방법을 택하면 더 확실합니다.

3.3 실제 글 분석과 학습 입력 예시

이제 실제로 사용자의 글을 한 편 가져와 어떻게 분석하고 구조화하는지 예를 들어 보겠습니다. 예를 들어 사용자 글이 다음과 같다고 가정합니다:

안녕하세요! 저는 개발자 홍길동입니다. 대학 시절부터 코딩에 흥미를 느껴 여러 프로젝트에 참여했어요. 특히 문제를 해결할 때 느끼는 성취감이 제 진로를 결정짓는 큰 동기가 되었습니다. 현재는 스타트업에서 일하며 매일 새로운 도전에 즐겁게 임하고 있습니다.

이 글을 분석하면 다음과 같은 문체 특징을 추출할 수 있습니다:

- 인사말에 “안녕하세요!”를 사용 -> **친근한 도입, 가벼운 느낌의 인사**
- “개발자 홍길동입니다.”로 본인 소개 -> **1인칭 자기소개, 존댓말 (입니다)**
- 전체적으로 반말과 존댓말 섞였음 (“참여했어요.” 평어체 + “입니다.” 존댓말) -> **약간 캐주얼하지만 기본적으로 공손한 어조**
- 이모티콘이나 과장된 표현 없음 -> **담백한 톤**
- 긍정적 내용(성취감, 즐겁게 임하고 있다) -> **낙관적/긍정 어조**

- 문장 길이는 대부분 20~30자 내외, 비교적 짧음 -> **간결한 편**
- 논리 전개: 자기 소개 -> 과거 -> 현재 순, **시간 순 나열 + 결론(진로 결정)**. 귀납적으로 사례(흥미, 성취감) 제시 후 결론(진로 결정 동기) 제시 -> **귀납+경험담형 전개**
- 자료 인용 없음, 개인 경험만 서술 -> **인용 스타일 해당 없음**
- 금지 요소: 특별히 없음 (비속어나 부적절 언어 없음) -> **클린함**

이를 바탕으로 사용자 홍길동의 문체 프로필을 만들어보면:

```
{
  "author_id": "hong_gildong",
  "style_profile": {
    "base_style": "간결체",
    "formality": "약간 격식 (혼합체)",
    "perspective": "1인칭 자기소개",
    "vocabulary_level": "일상어",
    "tone": ["친근함", "긍정적", "담담함"],
    "logic_style": ["경험담 나열", "귀납적"],
    "sentence_length": "중간 (20~30자 위주)",
    "connectives_usage": "보통",
    "citation_style": "N/A (개인경험)",
    "forbidden": []
  }
}
```

여기서 `formality`를 “약간 격식 (혼합체)”로 표현한 것은, 이 글에서 평서형과 존댓말 어미가 섞여 있어 반격식과 격식의 중간 톤임을 나타냅니다. `tone`에 “친근함”과 “긍정적”이 포함되어, 글 전반의 밝고 호의적인 분위기를 반영했습니다. `forbidden`은 별다른 금지 요소가 보이지 않아 빈 리스트로 두었습니다.

이런 프로필과 함께, 해당 사용자가 작성하기를 원하는 글의 **콘텐츠 요구사항**을 결합해 GPT에 입력하면 됩니다. 예를 들어 홍길동 씨의 문체를 학습한 GPT에게 “주어진 정보를 가지고 홍길동 문체로 블로그 글 작성해줘”라고 할 때, 내부적으로는 `<style_profile>...</style_profile><task>...</task>` 형식으로 프롬프트를 구성할 수 있습니다. fine-tuning을 한다면 아예 이런 형식의 데이터를 수집, 수백개 만들어 학습시킬 수 있습니다. 그렇게 하면 GPT 모델은 입력된 `style_profile`을 조건으로 삼아, 그에 맞는 어조와 형식으로 텍스트를 생성하게 됩니다.

OpenAI의 커스텀 GPT 기능에서도 유사한 접근이 가능합니다. 예컨대 **Custom Instructions**에 사용자의 문체 특징을 요약해서 적어둘 수 있습니다²⁷. “내 대답은 항상 짧은 단문으로, 격식을 약간 차리지만 친근한 톤으로 해줘.”와 같이 지침을 주는 겁니다. 더 나아가 OpenAI ChatGPT의 GPT 제작 도구에서는 **추가 지식**으로 사용자의 글 샘플 자체를 업로드할 수도 있습니다²⁷. 모델은 그 샘플을 참고하여 비슷한 스타일로 답변하게 됩니다. 즉, 구조화된 프로필을 명시적으로 주는 대신, **예시 대화나 문장**을 몇 개 제공하여 few-shot learning으로 문체를 따라하게 할 수도 있다는 뜻입니다. 다만 이 경우 모델이 일관되게 스타일을 따라할지 보장하기 어려울 수 있으므로, 가급적 명시적인 프로필+규칙을 함께 제시하는 것이 좋습니다.

정리하면, 문체 분석 결과를 **구조화된 메타데이터**로 표현하여 GPT에게 전달하면, 모델은 이를 조건으로 인식하고 글의 어조나 형식을 조절할 수 있습니다. JSON, YAML 등 형식은 도구적 수단일 뿐이며, 중요한 건 **문체 요소의 포괄적 기술**입니다. 또한 이러한 프로필을 만들기 위해서는 충분한 양의 사용자 글을 분석해 일반적인 경향성을 포착해야 하므로, **데이터 축적**도 필요합니다. 일단 프로필이 완성되고 나면, 커스텀 GPT는 어떤 글쓰기 상황이 주어지더라도 그 프로필을 참고해 **사용자 고유의 문체**로 출력을 내도록 설계될 수 있습니다.

4. 관련 자료 및 출처

커스텀 GPT를 구축하기 위해서는 OpenAI의 관련 문서와 NLP 연구 동향을 함께 참고하는 것이 좋습니다. 아래에 몇 가지 주요 자료와 연구를 정리합니다.

- **OpenAI 공식 문서 및 가이드:** OpenAI에서는 ChatGPT를 사용자가 원하는 방향으로 튜닝하기 위한 가이드라인을 제시하고 있습니다. 2023년 11월 발표된 “Introducing GPTs” 블로그에서는 사용자가 자체 GPT를 만들어 공유할 수 있는 기능을 소개하며, GPT를 만들 때 **세부 지침(Instructions)**과 **추가 지식(Knowledge)**을 넣어 원하는 행동을 유도할 수 있다고 설명합니다 ²⁷. 또한 OpenAI Help Center의 “Creating a GPT” 문서에서는 GPT를 생성하는 절차와 설정 항목을 안내하는데, 여기서 **GPT의 설명(Description)**과 **기능 제한(avoid behaviors)** 등을 설정할 수 있어 문체 가이드라인을 넣는 데 활용 가능합니다 ²⁷. OpenAI는 커스텀 GPT에서 사용자 데이터 업로드를 허용하여, 사용자의 글 모음 자체를 모델이 참고하게 할 수 있게 했습니다 ²⁸. 이를 통해 모델이 사용자 고유의 어휘와 스타일을 **지식 기반**으로 삼아 답변하게 하는 것입니다. OpenAI의 Fine-tuning 가이드도 참고할 만한데, GPT-3.5 등을 사용자 지정 스타일로 파인튜닝하는 방법이 문서화되어 있습니다. 이 경우 대규모 데이터와 비용이 들지만, 특정 기업의 **브랜드 톤**이나 **일관된 말투**를 학습시켜 응답에 반영하도록 할 수 있습니다. (OpenAI Fine-tuning 사용 시에는 대화 형식이 아닌 completion 형식으로 학습하는 제약 등이 있으므로, 최신 가이드를 확인해야 합니다.)
- **문체 분석 및 스타일 변환 연구:** 자연어 처리 분야에서 **텍스트 스타일 변환(Text Style Transfer)**은 활발한 연구 영역입니다. 스타일 변환은 주어진 문장의 내용을 유지하면서 선택된 문체 속성을 변경하는 것을 목표로 합니다 ²⁹ ³⁰. 예컨대 **감정(sentiment)**을 바꾸거나, **격식(formality)**을 바꾸거나, 현대어를 고어체로 바꾸는 작업 등이 포함됩니다 ³¹. 이러한 연구를 통해 축적된 방법론은 커스텀 GPT에도 응용될 수 있습니다. 특히 **대규모 언어모델(LLM)**을 활용한 스타일 변환 실험들이 주목할 만합니다. 2024년 Mukherjee 등은 대형 언어 모델들이 실제 스타일 변환에 얼마나 유용한지 분석한 연구에서, “거대 언어모델을 프롬프트만으로 활용할 경우 영어 외 언어(힌디어, 벵골어 등)에서는 평균 수준 성능에 그쳤지만, 소량 파인튜닝을 하면 성능이 크게 개선되어 SOTA에 근접했다”고 보고합니다 ³². 이는 한국어와 같은 비영어권 언어의 미묘한 문체 변환에도 **전용 데이터**와 **미세 조정**이 필요함을 시사합니다. 다시 말해, 커스텀 GPT가 한국어 사용자 문체를 충실히 모방하려면 몇 가지 예시만으로는 한계가 있고, 가능한 한 사용자 문체 데이터로 모델을 추가 학습시키는 것이 효과적일 수 있습니다.
- **한국어 문체 분석/변환 사례:** 한국어에서도 스타일 변환 연구가 진행되어 왔습니다. 예를 들면, 한국어 챗봇 대화에서 **채팅체(인터넷 슬랭, 반말)**를 **문어체(격식체)**로 바꾸는 연구가 있었습니다 ³³. 2023년 정영준 등은 비지도 학습으로 채팅체를 문어체로 자동 변환하는 모델을 제안하였는데, 이 모델은 기계번역 기법을 응용하여 스타일이 다른 두 말뭉치를 서로 변환하도록 학습했습니다 ³⁰. 이러한 접근으로 채팅체와 문어체 사이의 **단어 대응 사전**도 구축할 수 있었다고 합니다 ³⁰. 이는 문체 변환을 위해 **평행 데이터 없이도** 스타일 차이를 학습하는 방법을 보여주며, 커스텀 GPT에도 시사하는 바가 큼니다 (예: 사용자 문체와 일반 문체 간 차이를 학습시켜 자유자재로 전환). 또 다른 예로, 국내 NLP 기업이나 연구팀에서 **인공지능 맞춤 글쓰기 보조**를 위해 문체 변환 기술을 활용하고 있습니다. 한 사례로 언급되는 **TextNet 서비스**에서는 “MBTI와 매핑한 문체 제공을 통해 페르소나와 일치하는 코퍼스 생성 가능”한 **문체 생성 및 변환 서비스**를 제공했다고 합니다 ³⁴. 이는 이용자의 성격 유형에 어울리는 문체로 텍스트를 바꿔주는 기능으로 추측되는데, 실제 적용 분야가 챗봇, 커머스 상품 설명, 고객센터 응대 등 광범위했습니다 ³⁵. 이러한 실용 사례는 커스텀 GPT가 사용자 개인의 문체뿐만 아니라, **원하는 페르소나나 브랜드 스타일**에 맞게 조정될 수 있음을 보여줍니다.
- **논문 및 참고 자료:** 문체와 관련하여 보다 폭넓은 이해를 위해 참고할 만한 문헌으로는, 전통적으로 **스타일로메트리(Stylometry)** 분야의 연구들이 있습니다. 이것은 문체를 정량적으로 분석하여 필자를 식별하거나 텍스트를 분류하는 분야로, 사용자 문체 프로파일링에 도움이 될 수 있습니다. 또한 Pennington et al.의 글쓰기 가이드나 한국어 **글쓰기 교육 자료**(예: 서울대학교 글쓰기 센터의 자료 ³⁶)에서도 문체 요소에 대한 통찰을 얻을 수 있습니다. 커스텀 GPT를 만들 때는 이러한 인문학적/언어학적 통찰과 AI 기술을 결합하는 것이 중요합니다.

마지막으로, OpenAI 커뮤니티 포럼 등에는 **사용자들이 GPT를 특정 문체로 활용하는** 사례와 팁이 공유되고 있습니다. 예를 들어 한 사용자는 짧은 공문 메모 작성을 도와줄 GPT를 만들면서 “격식 있고 간결한 공지 문체”를 GPT에게 학습 시키는 방법을 논의하기도 했습니다 (여기서 권장된 방법은 해당 스타일의 예문을 여러 개 GPT에 주어 few-shot 세팅을 하는 것이었습니다). 이러한 커뮤니티 지식도 실무적으로 많은 도움을 줄 수 있습니다.

참고문헌 및 출처 요약: 한국민족문화대백과의 문체 항목 ¹ 을 통해 문체 유형의 전통적인 분류를 참조했고, 브런치스 토리 등의 글쓰기 조언글 ¹³ ¹¹ 로부터 SNS/보고서 문체 실무 팁을 인용했습니다. 또한 문체 변환 연구 관련으로 ACL 등에서 나온 논문들의 개요 ²⁹ ³² 와 국내 학술대회의 논문 초록 ³⁰ 등을 참고했습니다. OpenAI 관련 내용은 공식 도움말 ²⁷ 을 기반으로 했습니다. 이러한 자료들은 커스텀 GPT 구현 시 문체를 체계적으로 다루는 데 근거로 삼을 수 있을 것입니다.

以上の 내용과 출처들을 종합하면, 사용자 한국어 글의 문체를 다각도로 분석하고 이를 구조화하여 GPT 모델에 적용함으로써, 다양한 글 목적과 어조에 유연하게 대응하는 **맞춤형 GPT**를 구축할 수 있습니다. 이 과정에서 언어적 통찰과 기술적 수단을 함께 활용하는 것이 성공의 열쇠입니다.

¹ 문체(文體) - 한국민족문화대백과사전

<https://encykorea.aks.ac.kr/Article/E0019695>

² ¹⁹ '-함', '-음'의 개조식 문장, 천황 시대의 권위주의 잔재 - 오마이뉴스

https://www.ohmynews.com/NWS_Web/View/at_pg.aspx?CNTN_CD=A0002343769

³ ⁴ ⁵ ⁶ ¹⁷ 문어체와 구어체의 차이

<https://kimsway.tistory.com/117>

⁷ Inductive vs. Deductive Writing - Purdue University Global Academic Success Center

<https://purdueglobalwriting.center/2015/02/25/inductive-vs-deductive-writing/>

⁸ 논문 인용과 참고문헌 표기법(APA/MLA 스타일), "et al." 뜻

<https://blog.essayreview.co.kr/academic/reference-citation-et-al/>

⁹ ¹⁰ ¹¹ ¹² ¹³ ¹⁴ SNS 웹문서 글쓰기

<https://brunch.co.kr/@bABU/73>

¹⁵ ¹⁶ 연설문, 어떻게 써야 할까

<https://brunch.co.kr/@2rDA/158>

¹⁸ 보고서 개조식 문장 쓰기, 강신정 강사님 : 네이버 블로그

<https://blog.naver.com/alsemffp234/222387884486>

²⁰ ²¹ 서술형과 개조식의 장점을 섞은 ‘서술형 개조식’

<https://brunch.co.kr/@itrendlab/29>

²² ²³ ²⁴ ²⁵ ²⁶ 자기소개서 < 취업/진학 | 하남경영고등학교

<https://hanambiz-h.goegh.kr/hanambiz-h/cm/cntnts/cntntsView.do?mi=4299&cntntsId=782>

²⁷ ²⁸ Creating a GPT | OpenAI Help Center

<https://help.openai.com/en/articles/8554397-creating-a-gpt>

²⁹ ³¹ Text Style Transfer | Papers With Code

<https://paperswithcode.com/task/text-style-transfer/codeless>

³⁰ ³³ 비지도 기계 번역을 이용한 채팅체 문체 변환

<https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artId=ART002925540>

32 [2406.05885] Are Large Language Models Actually Good at Text Style Transfer?

<https://arxiv.org/abs/2406.05885>

34 35 yis.yonsei.ac.kr

<https://yis.yonsei.ac.kr/korean/jobs01.do?mode=download&articleNo=131214&attachNo=111676>

36 글쓰기 마무리하기 - 인용하기와 인용출처 표기 - 서울대학교 온라인 ...

<https://owl.snu.ac.kr/2430/>