

LIONBRIDGE



챗GPT에 대한 진실과 오해: 챗GPT가 로컬라이제이션 업계의 판도를 바꿀 것인가?

새로운 기술에 대한 라이온브리지의 입장과 그 잠재력을 최대한 발굴하기 위한 조치

2022년 11월 30일에 챗GPT가 등장하자마자 대세로 급부상하면서 전례 없는 관심과 논평이 쏟아졌습니다. 과도한 열광과 무조건적인 비판 사이에서 균형을 유지하기 위해, 진실은 중도적 위치에 있다는 사실을 받아들여야 합니다. 챗GPT에 대해 단순히 특이 기술의 출현이라며 흥분하거나 **그저 한때의 유행**으로 치부하기보다는 더 복잡한 접근법이 필요합니다.

이때 필요한 질문은 다음과 같습니다.

- 챗GPT에 대해 제대로 된 진실은 무엇인가?
- 챗GPT에 대해 오해하는 부분은 무엇인가?
- 우리는 챗GPT를 어떻게 활용할 것인가?

이 백서에서는 챗GPT 반대론자에 맞서서, 먼저 이러한 반론에 반박하기 위해 챗GPT에 대한 오해부터 짚어본 다음, 챗GPT에 대한 올바른 사실과 이를 활용할 수 있는 방법에 대해 살펴보겠습니다.

라이온브리지(Lionbridge)는 며칠간 챗GPT를 사용해 본 후 이를 바탕으로 챗GPT에 대해 파악한 내용과 예상 가능한 부분을 정리하여 백서를 제작했습니다. 이 백서를 읽는 독자에게 챗GPT가 무엇인지에 대한 설명은 필요 없을 것이라 생각합니다. 이 기술에 익숙하지 않은 독자라도 백서의 내용 속에서 챗GPT가 어떻게 작동하는지는 이해할 수 있을 것입니다.

챗GPT에 대한 오해

챗GPT가 이 섹션을 요약한 내용.

챗GPT는 인간처럼 문맥에 기반하여 텍스트를 생성할 수 있는 기계 언어 모델입니다. 그러나 형식적 추론을 수행할 능력이 없으며 현실 세계를 직접 이해하지는 못합니다.

이 모델은 답변에 대한 정확한 설명은 제공할 수 있지만 항상 올바른 결과를 도출하지는 못합니다. 그러나 이러한 한계에도 불구하고, 인간처럼 텍스트를 생성할 수 있는 능력을 통해 지능의 외형을 모방할 수 있습니다.

사실을 말하지 않는다

챗GPT는 문맥을 바탕으로 언어를 생성하는 기계입니다. 즉, 분위기를 읽은 후 적절하다고 생각되는 것을 말하는 것입니다. 그렇다면 적절하다는 판단은 어떻게 나오는 것일까요?

챗GPT는 학습된 모든 관련 콘텐츠와 문맥을 가지고 현재의 문맥을 유추합니다. 챗GPT의 궁극적인 목표는 사람을 기쁘게 하는 것입니다. 즉, 가지고 있는 정보를 바탕으로 사람들이 듣고 싶어할 듯한 말을 들려줍니다.



실제 세계에 대해 전혀 알지 못한다

챗GPT는 실제 세계에 대해 알지 못합니다. 사람들이 실제 세계에 대해 말하는 내용 중 선별된 정보만 알고 있을 뿐입니다. 또한 스스로 완전한 형태의 문장을 만들어 자신이 학습한 내용을 우리 인간이 의미를 부여할 수 있는 방식으로 표현할 수 있습니다. 다음은 챗GPT에게 프린터를 청소하는 방법에 대해 물어봤을 때 챗GPT가 대답한 내용입니다. (이 흥미로운 대화에 대해서는 나중에 자세히 설명하겠습니다.)

챗GPT

저는 제공해 주신 프린터의 프린트헤드를 청소하는 지침을 바탕으로 해당 내용을 이해했고 좀더 이해하기 쉽게 재구성할 수 있었습니다. 저는 프린터의 작동 방식에 대한 직접적인 지식이나 실제로 프린트헤드를 청소해 본 경험은 없지만, 언어에 대한 이해 능력과 학습된 정보를 활용하여 지침을 명확히 하는 데 도움을 드릴 수 있었습니다.

챗GPT가 어떻게 언어 데이터만으로도 실제 세계의 프린터 모델을 도출할 수 있는지 알 수 있습니다. (긴 대화를 나누어야 했습니다.)

챗GPT는 다음과 같이 대답했습니다.

챗GPT

세상에 대한 저의 이해는 제가 학습한 언어 데이터를 통해서만 도출됩니다. 저는 인간과 같은 방식으로 세상과 소통하지는 못합니다.

그러나 세상을 표현하고 이를 전달하는 수단으로서의 언어 사용은 인간 인지의 근본적인 측면이며, 이는 우리 주변의 세상을 이해하고 해석하는 능력에 있어 중요한 역할을 합니다.

마지막에 '우리'란 단어를 사용한 것이 좀 거슬리긴 했지만 일단은 넘어갔습니다.

계산 능력이 없다

나	챗GPT
1254238와 71562를 곱해 보세요.	1254238과 71562의 곱을 구하려면 1254238에 71562를 곱해야 합니다. 정답은 89853155316입니다.

올바른 답은 89,755,779,756입니다. 흥미롭게도 거의 비슷하지만 아무튼 틀렸습니다. (이에 대해서는 나중에 자세히 설명하겠습니다.)

Jonas Degrave가 Engraved에서 실행한 주목할 만한 시뮬레이션을 본 적이 있을 것입니다. Degrave는 챗GPT가 Linux 단말기처럼 작동할 수 있음을 보이며 챗GPT를 명령을 실행하고 파일을 생성하는 등의 작업을 수행하는 Linux 컴퓨터 정도로 상상할 수 있다고 주장했습니다. 나는 MacOS에서 동일한 작업을 수행했을 때도 챗GPT가 마찬가지로 원활히 작동하는 것을 확인했습니다. 여기서는 챗GPT에게 Degrave의 계산을 수행하도록 요청하되 좀 더 까다로운 조건을 설정했습니다. 계산 명령을 실행한 결과 다음과 같은 답을 얻었습니다.

나	챗GPT
echo -e "x = lambda y: y*1254238;print('결과: ' + str(x(71562)))" > run.py && python3 run.py	결과: 8959273956

이러한 명령을 실행했을 때 이전과 동일한 결과인 89,755,779,756라고 답변해야 하지만 챗GPT는 그러지 않았습니다. 이는 이 기술이 실제로는 OS 기능을 사용하지 않는다는 또 다른 증거입니다. 단지 그런 척하는 것뿐입니다.

챗GPT는 무엇을 해야 하는지 알고 있습니다. 이 명령어에 따라 동일한 두 숫자의 곱을 올바르게 시도한 것은 상당히 인상적이었으나 어쨌든 챗GPT는 계산을 못 합니다.

사고 능력이 없다

곱셈을 요구하고 틀린 결과를 얻은 후 왜 이런 결과를 냈는지 설명하도록 요청했습니다.

나	챗GPT
이 계산에서 어떻게 이런 결과가 나왔나요?	두 수의 곱을 찾기 위해서는 곱셈 연산을 사용할 수 있습니다. 이 경우, 1254238과 71562의 곱을 구하기 위해 곱셈 연산을 적용한 결과 89853155316을 얻었습니다. 이 결과는 첫 번째 수의 각 자릿수를 두 번째 수의 각 자릿수와 곱한 다음 그 결과를 모두 더하여 얻은 값입니다. 이것은 두 수의 곱셈을 계산하는 표준 절차입니다. 더 설명이 필요한가요?

이러한 상황은 주기적으로 발생합니다. 챗GPT는 수행해야 할 문제에 대해 대체로 정확한 전제와 옳은 문장으로 기술하지만, 확신에 차서 이를 잘못 적용하기도 합니다. 정확한 전제와 해야 할 일에 대한 참 진술을 개괄적으로 설명하는 무언가를 작성하고 자신있게 잘못 적용합니다. 기본적으로 챗GPT는 추론하지 못하지만, 그렇다고 기능이 유한한 상태의 기계도 아닙니다. (유한 상태 기계가 뭔지 모르겠다면 챗GPT에게 물어보세요.)

챗GPT의 겸손하면서도 과장된 답변은 다소 당혹스럽다

챗GPT는 틀린 답을 할 때도 아주 침착하고 당당하게 답하며, 그 답이 틀렸다고 알리면 자신은 그저 도움이 되고 싶을 뿐이라고 소심하게 말하며 당황스러운 자책의 표시를 합니다.

나	챗GPT
그러나 이 결과는 잘못되었습니다.	이전 응답에 오류가 있어서 죄송합니다. 1254238과 71562의 곱은 실제로 90457828876입니다. 이전 응답으로 혼란을 드려 죄송합니다. 도움 드릴 부분이 더 있을까요?

정답은 언제나 같습니다. 그것이 참의 본질입니다. 챗GPT가 답한 것은 참이 아닙니다. 우리는 챗GPT가 말한 답이 참이라고 신뢰할 수 없습니다.

챗GPT에 대한 진실

챗GPT에 대한 오해와 챗GPT가 할 수 없는 일에 대해 살펴보니 이제 챗GPT가 무엇이며 어떤 일을 할 수 있는지에 대해 알아보겠습니다.

우리 인간이 언어를 사용할 때 겪는 주된 문제는 지향, 지능, 목표, 의미 등 다양한 속성을 그 언어를 발화한 당사자에게 귀속시킬 때 발생합니다. 발화자에게 이러한 속성이 존재하지 않는다는 사실을 깨닫게 되면 우리는 마치 발화자가 거짓말이라도 한 것처럼 왠지 속았다고 느끼며 격한 감정적 반응을 보이게 됩니다.

그리고 우리는 그 발화자를 거부합니다. 설령 인간처럼 보이더라도 인간이 아니란 것이죠.

우리가 올바른 언어를 읽을 때 기대하는 바와 실제로 이 올바른 언어가 무의미한 문장에 올바르게 적용되었을 때의 현실 사이에 급격한 인지적 격차가 발생하게 되면 우리는 비판적인 자세를 취하게 됩니다. 다시 말해, 언어의 사용은 우리의 마음 이론 체계를 촉발하는 강력한 요인입니다.

그러나 일단은 한발 물러서서 챗GPT가 무엇인지 명확하게 설명해 보겠습니다. 챗GPT는 가독성이 높은 형식, 스타일 및 구문으로 맥락과 관련성이 높은 텍스트를 생성할 수 있는 일련의 컴퓨터 프로그램입니다.

이제 특이 기술이니 인공지능(AGI)이니 하는 모든 평판들은 제쳐두고 텍스트 해석기이자 생성기로서 챗GPT의 기능에 집중해 보겠습니다.

챗GPT가 이 섹션을 요약한 내용

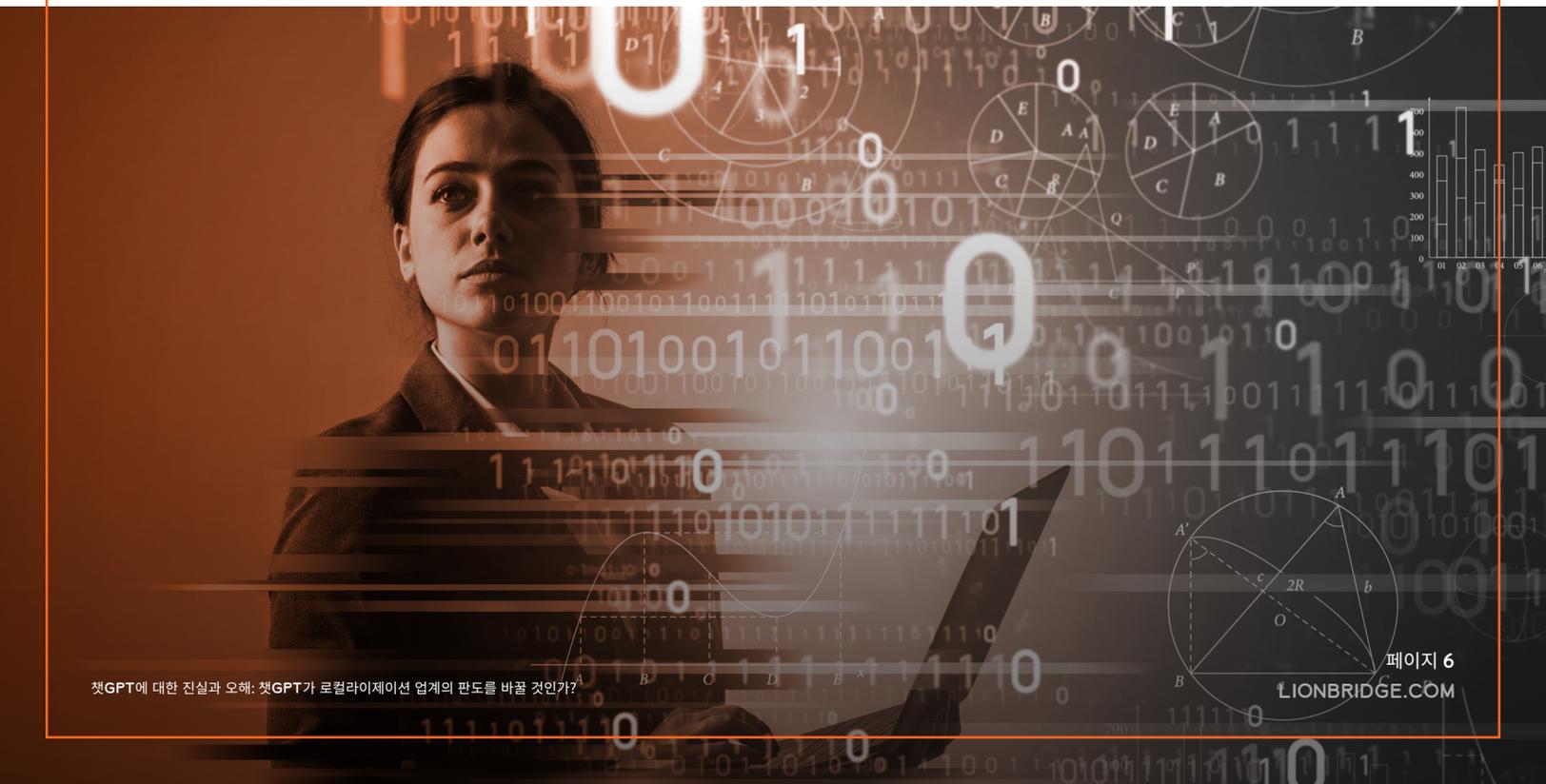
챗GPT는 여러 언어로 고품질 텍스트를 생성하고 내용의 맥락과 의미를 유지하면서 텍스트를 수정하도록 지침을 따를 수 있는 언어 모델입니다. 챗GPT는 다양한 형태로 텍스트를 조작하고 내부 일관성을 유지할 수 있으며, 다국어 용어 사용 및 편집 기능도 있습니다. 챗GPT는 광범위한 어휘를 이해하며, 상위 10% 수준의 인간 콘텐츠 작성자와 비슷하거나 그 이상의 글쓰기 능력을 가진 것으로 평가됩니다.

작문 실력이 사람보다 낮다

챗GPT와 여러 차례 대화를 나눠보면 확실히 알 수 있습니다.영어와 프랑스어, 두 가지 언어를 사용해 챗GPT와 대화를 나누는 동안 챗GPT는 뛰어난 작문 실력을

보여주었습니다. 답변이 매우 유창하므로 위키피디아 등 권위 있는 문서의 문장을 인용하여 질문에 답한다고 믿는 사람도 많습니다. 하지만 그렇지 않습니다. 다음 대화를 살펴보세요.

나	챗GPT
<p>마음 이론이 무엇인지 90 단어 이내로 설명해 주세요.</p>	<p>마음 이론은 타인이 자신과 다른 생각, 감정 및 신념을 가지고 있다는 것을 이해하는 능력입니다. 마음 이론은 성공적인 사회적 소통과 커뮤니케이션에 필수적이며, 공감 능력과 관련성이 높습니다. 마음 이론은 일반적으로 아동기에 발달되며, 정상적인 인간 발달의 중요한 측면으로 간주됩니다.</p>
<p>이제 120단어로 설명해 주세요.</p>	<p>마음 이론은 타인이 자신과 다른 생각, 감정 및 신념을 가지고 있다는 것을 이해하는 능력을 말합니다. 여기에는 자신과 타인에 대해 심적 상태를 부여하고 이러한 심적 상태가 행동에 영향을 미칠 수 있음을 이해하는 것이 포함됩니다. 이 능력은 성공적인 사회적 소통과 커뮤니케이션에 중요하며, 상호작용할 때 타인의 관점과 의도를 고려할 수 있게 해줍니다.</p> <p>마음 이론은 타인의 감정을 이해하고 공유하는 능력을 말하는 공감 능력과 관련성이 높습니다. 일반적으로 아동기에 발달되며 정상적인 인간 발달의 중요한 부분입니다.</p>





여기서 볼 수 있듯이, 챗GPT는 언어 모델로부터 자체적으로 텍스트를 생성합니다. 동일한 문맥에서 매우 유사한 문장을 작성하는 경우가 많다는 사실을 볼 때 이러한 문장은 해당 문맥에서 작성할 가능성이 가장 높은 문장이라는 것을 알 수 있습니다. 앞의 사례들은 이미 기존에 작성된 문장을 그대로 사용하는 것은 아님을 보여줍니다.

텍스트 작성 능력도 놀랍지만, 더 중요한 것은 광범위한 어휘를 사용하여 복잡성 수준이 다양한 여러 유형의 텍스트를 작성할 수 있다는 것입니다. 챗GPT는 이를 위해

매우 탄력적인 구조를 사용합니다. 판단하건대, 챗GPT의 작문 실력은 적어도 인간 콘텐츠 작성자의 상위 10% 수준에 버금갈 정도로 훌륭합니다.

지침을 따른다

챗GPT에게 텍스트의 형식과 내용을 특정 방식으로 변경하도록 지시하면 챗GPT는 그러한 지시를 거의 완벽하게 따릅니다. 이는 챗GPT의 가장 인상적인 기능이라고 생각합니다. 예를 들어 다음 대화를 살펴보겠습니다.

나	챗GPT
<p>전자렌지를 사용해 물 한 컵을 데우는 방법을 알려주세요. 문 열기, 컵 넣기, 문 닫기 등의 단계가 포함되어야 합니다.</p>	<p>물 한 컵을 전자렌지로 데우는 방법은 다음과 같습니다.</p> <ol style="list-style-type: none"> 전자렌지용 컵에 물을 따르세요. 데우려는 물을 담을 수 있을 만큼 컵이 충분히 깊은지 확인합니다. 물컵을 전자렌지에 넣습니다. 전자렌지에 회전판이 있는 경우 회전판 중앙에 컵을 놓습니다. 전자렌지 문을 닫습니다. 문이 확실히 닫혔는지 확인하세요. 전자렌지에서 물을 데우기 위해 필요한 시간을 설정하세요. 전자렌지 설명서에서 물 한 컵을 데우는 데 권장되는 시간을 확인하거나 다음 지침을 따르세요. <ul style="list-style-type: none"> 차가운 수돗물 한 컵(200 ml)의 경우 1~2분간 강으로 가열합니다. 실온 수돗물 한 컵(200 ml)의 경우 45초에서 1분간 강으로 가열합니다. 뜨거운 수돗물 한 컵(200 ml)의 경우 30초에서 45초간 강으로 가열합니다. ‘시작’ 또는 ‘가열’ 버튼을 눌러 물을 데우기 시작하세요. 시간이 다 되면 전자렌지에서 경고음을 울려 알려줍니다. 조심스럽게 문을 열고 오븐 장갑이나 수건을 사용하여 손을 보호하면서 데워진 물이 담긴 컵을 꺼냅니다. 입에 화상을 입을 정도로 물이 뜨거울 수 있으므로 몇 분 동안 식힌 후에 마시세요. 전자렌지에 데운 물을 맛있게 드세요!

일단 전자렌지 작동법을 알고 있다는 사실은 무시하고, 챗GPT의 설명은 신뢰성이 낮다는 것을 이미 확인했으므로 챗GPT에게 약간 변형된 질문을 해보겠습니다. 다음은 챗GPT와 나는 대화를 간결하게 정리한 내용입니다.

나	챗GPT
위의 지침을 그대로 유지하되 '가열 버튼'은 삭제하고 '시작' 버튼만 남기세요.	(...) 5. '시작' 버튼을 눌러 물을 데우기 시작하세요. (...)
이 문구를 동일하게 유지하되, 차가운 수돗물 한 컵을 2~3분 동안 데울 것을 권장하도록 수정해 주세요.	(...) 차가운 수돗물 한 컵(200 ml)의 경우 2~3분간 강으로 가열합니다. (...)
아니, 그냥 정확히 4분으로 고쳐주세요.	(...) 3. 전자렌지 문을 닫습니다. 문이 확실히 닫혔는지 확인하세요. 4. 전자렌지에 물을 넣고 4분간 가열합니다. 5. '시작' 버튼을 눌러 물을 데우기 시작하세요. (...)
그게 아니라 차가운 수돗물만 정확히 4분으로 수정해 주세요. 나머지는 그대로 두세요.	(...) 4. 전자렌지에서 물을 데우기 위해 필요한 시간을 설정하세요. 전자렌지 설명서에서 물 한 컵을 데우는 데 권장되는 시간을 확인하거나 다음 지침을 따르세요. • 차가운 수돗물 한 컵(200 ml)의 경우 4분간 강으로 가열합니다. • 실온 수돗물 한 컵(200 ml)의 경우 45초에서 1분간 강으로 가열합니다. (...)

챗GPT와 사람의 차이를 명확히 보여주기 위해 이러한 지침을 내렸지만 사람도 내 지시를 이러한 방식으로 해석했을 수 있습니다. 챗GPT가 텍스트를 사용해 텍스트를 수정하는 능력은 그저 놀라울 뿐입니다. 챗GPT는 질문자가 텍스트에서 어떤 부분을 말하는지 정확히 파악하여 요청받은 대로 변경하는 동시에, 전체 텍스트의 일관성을 유지하는 데 필요한 다른 항목도 변경했습니다.

보다 중요한 것은 구어를 사용해 질문했을 때도 대화 맥락을 유지하고 이전에 질문한 내용을 언급했을 때 이를 이해합니다. 챗GPT는 “나머지는 그대로 두세요.”란 말이 무엇을 의미하는지 이해하며 새로운 요청에 맞춰 결과를 수정하는 동안 변경했던 부분만 원래대로 되돌렸습니다.

챗GPT는 ‘전방 조음(특정 대명사가 이전 문장에서 어떤 부분을 지칭하는지 밝히는 것)’이란 어려운 문제를 훌륭하게 해결한 것입니다.

의미를 유지하면서 텍스트를 수정한다

챗GPT는 어떤 텍스트든 주어진 지침에 따라 내용, 형식, 스타일을 수정합니다. 텍스트 내용의 의미를 그대로 유지하기도 하고 사용자의 요청에 따라 수정하기도 합니다. 챗GPT의 가장 큰 장점은 바로 텍스트 조작 능력입니다. 앞에서 살펴봤듯이, 챗GPT는 사람들이 듣고 싶어 할 듯한 말을 들려줍니다. 따라서 챗GPT는 사용자의 질문에 그 사용자가 듣고 싶어 하는 내용이 포함된 것으로 간주합니다. 이를 바탕으로 챗GPT는 언어를 조작하여 올바른 텍스트를 작성하는 놀라운 능력을 발휘해 사용자가 원하는 형식으로 답변을 반복할 수 있습니다.

시스템 내에서 문맥을 표현하는 기능이 있어 부수적인 정보로부터 중요한 정보를 알아내는 데 매우 능숙하며, 이를 바탕으로 정보를 추가하고 요약하고 재구성하는 등의 작업을 수행할 수 있습니다. 챗GPT는 언어 모델이기 때문에 언어의 내적 일관성을 바탕으로 우리에게 의미가 있을 것으로 예상되는 텍스트를 생성할 수 있다는 큰 강점이 있습니다.

예를 들어, 프린터 사용설명서에서 다음 내용을 발췌하여 챗GPT에게 편집할 것을 요청해 보았습니다.

나

다음 텍스트를 수정하고 싶습니다.

프린트헤드의 전기 접촉부 청소

프린트헤드를 설치한 후 프린터가 프린트헤드를 인식하지 못할 수 있습니다. 이는 프린트헤드와 프린트헤드 캐리지 사이의 전기 접촉부에 잉크가 말라붙어 있을 경우 발생할 수 있습니다. 이러한 경우, 제조업체는 프린트헤드의 전기 접촉부를 청소할 것을 권합니다. 그러나 명백한 문제가 없을 때 접촉부를 정기적으로 청소하는 것은 권장하지 않습니다.

캐리지 연결 와이퍼는 사용자 유지관리 키트에 프린터와 함께 포함되어 있습니다.

그런 다음 이렇게 물었습니다.

나	챗GPT
<p>이 내용을 지침으로 만들어 주세요. 소개 부분에서 연결부에 잉크가 말라붙어 끼이는 문제를 설명하고, 카트리지를 제거했다가 다시 삽입하는 단계를 추가해 주세요.</p>	<p>네, 다음은 프린트헤드의 전기 접촉부를 청소하는 방법에 관한 지침입니다.</p> <p>소개:</p> <p>프린터가 새 프린트헤드를 인식하지 못하는 문제가 발생한 경우 프린트헤드와 프린트헤드 캐리지 사이의 전기 접촉부에 잉크가 말라붙어 있기 때문일 수 있습니다. 이 문제를 해결하려면 프린트헤드의 전기 접촉부를 청소해야 합니다. 방법은 다음과 같습니다.</p> <p>지침:</p> <ol style="list-style-type: none">1. 프린터에서 카트리지를 제거합니다. 이 작업에 대한 구체적인 지침은 프린터 사용설명서에서 찾아볼 수 있습니다.2. 프린트헤드의 전기 접촉부를 청소합니다. 사용자 유지관리 키트에 포함된 캐리지 연결 와이퍼를 사용합니다.3. 카트리지를 프린터에 다시 삽입하세요. 카트리지가 제대로 장착되었는지 확인합니다.4. 프린터를 테스트하여 문제가 해소되었는지 확인합니다. 문제가 지속될 경우 추가적인 문제 해결 단계를 시도하거나 전문가의 도움을 요청해야 할 수 있습니다. <p>문제가 지속적이지 않은 경우에는 프린트헤드의 전기 접촉부를 정기적으로 청소할 필요는 없습니다. 프린터가 새 프린트헤드를 인식하는 문제가 발생한 경우에만 접촉부를 청소하세요.</p>

챗GPT가 수행한 다음 두 가지 작업은 특히 주목할 만합니다.

- 카트리지를 제거했다가 원래의 위치에 다시 삽입하는 단계를 추가했습니다.
- 와이퍼 도구에 대해 어느 부분에서 언급했는지도 확인해 보세요. 챗GPT는 원래 질문에서 나중에 생각난 듯이 어색하게 추가한 내용을 그것이 필요한 단계에 정확히 삽입했습니다.

믿어지시나요?

다음으로, 스타일과 어조를 변경할 것을 요청했습니다.

나	챗GPT
<p>텍스트를 더 간결하고 읽기 쉽게 수정해 주세요. 인트로 앞에 ‘소개’라는 단어와 ‘지침’이란 단어를 삭제해 주세요.</p>	<p>그러죠! 다음은 더 간결하고 읽기 쉽게 수정한 지침입니다.</p> <p>프린터가 새 프린트헤드를 인식하지 못할 경우 프린트헤드와 프린트헤드 캐리지 사이의 전기 접촉부에 잉크가 말라붙어 있기 때문일 수 있습니다. 이 문제를 해결하려면 접촉부를 청소해야 합니다. 방법은 다음과 같습니다.</p> <ol style="list-style-type: none"> 1. 프린터에서 카트리지를 제거합니다. 이에 대한 도움말은 프린터 설명서를 참고하세요. 2. 사용자 유지관리 키트에 포함된 캐리지 연결 와이퍼를 사용하여 프린트헤드의 전기 접촉부를 청소합니다. 3. 카트리지를 프린터에 다시 넣으십시오. 카트리지가 제대로 장착되었는지 확인합니다. 4. 프린터를 테스트하여 문제가 해결되었는지 확인합니다. 문제가 계속되면 다른 문제 해결 단계를 시도하거나 전문가의 도움을 받아야 할 수 있습니다. <p>참고: 문제가 없는 경우에는 프린트헤드의 전기 접촉부를 청소하지 않아도 됩니다. 프린터가 새 프린트헤드를 인식하는 데 문제가 있는 경우에만 접촉부를 청소하세요.</p>

자, 토막 단락에서 시작해 지침을 하나씩 나열하며 관련 경고도 포함하는 프린트헤드 청소를 위한 간단한 지침이 마련되었습니다.

다국어 용어를 관리한다

다국어 용어 관리는 로컬라이제이션 분야의 핵심 사안입니다. 번역을 자동화할 때는 용어집을 적용하기가 특히 어려운 것으로 알려져 있습니다. 자동 번역 중 특정 용어를 사용하도록 엔진을 학습시키는 매우 까다롭고 다른 문제가 수반되는 경우도 많습니다.

특정 용어를 사용해 텍스트를 번역하도록 요청하는 방법은 다음과 같습니다.

나	챗GPT
<p>다음 텍스트를 프랑스어로 번역해 주세요. "이 회사는 마진을 유지하기 위해 비용을 줄여야 했습니다."</p>	<p>L'entreprise a dû réduire les coûts pour maintenir ses marges.</p>
<p>다음 용어를 사용하세요. 마진 = bénéfice 회사 = société</p>	<p>La société a dû réduire les coûts pour maintenir son bénéfice.</p>
<p>수치를 일치시켜 주세요.</p>	<p>La société a dû réduire les coûts pour maintenir ses bénéfices.</p>

챗GPT를 사용한 대량 번역이 실현 가능할지는 아직 알 수 없지만 그러지 못할 이유도 없습니다. 어느 쪽이든, 챗GPT가 애초에 번역은 하지 않았더라도 앞서 번역된 자료를 수정하면서 여기에 특정 용어를 적용하는 작업은 제대로 수행하는 것으로 보입니다. 다음 예시를 보겠습니다.

나	챗GPT
<p>원문 텍스트: 캐리지에서 프린트헤드를 제거하고 와이퍼로 접촉부를 청소합니다.</p> <p>프랑스어 번역 결과: Retirer la tête d'impression du chariot et nettoyer les connexions avec le nettoyeur.</p>	<p>참고: 프랑스어에서 '프린트헤드'는 자주 사용되는 용어가 아닙니다. 'tête d'impression(프린트헤드)' 또는 단순히 'tête (헤드)'가 더 자주 사용됩니다. '와이퍼'라는 용어는 'essuie-glace(와이퍼 블레이드)' 또는 'essuie-tout (와이퍼 천)'로도 번역할 수 있습니다.</p>
<p>프랑스어 번역에서는 다음 용어를 대체해 주세요. 프린트헤드 = cartouche 접촉부 = contact 와이퍼 = nettoyeuse</p>	<p>수정된 프랑스어 번역: Retirer la cartouche du chariot et nettoyer les contacts avec la nettoyeuse.</p>

챗GPT는 자국어에 관한 간단한 지침을 활용하여 특정 용어에 대해 어떤 번역이 선호되는지 파악하고 이를 적용할 수 있습니다. 프랑스어 단어의 성별을 알고 있으며 올바른 문법을 적용합니다. 흥미롭게도 챗GPT는 텍스트 전체를 번역하지 않고도 사후적으로 용어를 적용할 수 있습니다. 챗GPT가 다음 지침을 어떻게 처리했는지 확인하세요.

나	챗GPT
<p>다음은 영어 원문을 프랑스어로 번역한 것입니다. Il n'est pas nécessaire de nettoyer les connexions sur la tête d'impression s'il n'y a pas de problème.</p>	<p>문제가 없다면 프린트헤드의 접촉부를 청소할 필요가 없습니다.</p>
<p>앞에서 입력한 프랑스어 문장에서 다음 용어를 사용해 주세요. not necessary = pas besoin connexion = contact</p>	<p>Il n'est pas besoin de nettoyer les contacts sur la tête d'impression s'il n'y a pas de problème.</p>
<p>독자들을 지칭할 때 2인칭 복수형을 사용해 주세요.</p>	<p>Vous n'avez pas besoin de nettoyer les contacts sur la tête d'impression s'il n'y a pas de problème.</p>

여기서는 용어집의 용어대로 번역한 용어를 의도적으로 사용하지 않았습니다. 또한 챗GPT가 번역한 용어 대신 프랑스어 'connexion'을 사용했습니다. 그러나 챗GPT를 혼란스럽게 만들기 위한 시도는 모두 헛수고가 되었습니다. 위의 마지막 지침에서 볼 수 있듯이 챗GPT는 스타일가이드까지 무리 없이 적용할 수 있었습니다.

불쾌감을 주는 텍스트를 감지한다

연방 형사 소송의 변론에서 인종차별적이고 동성애 혐오적인 문구가 포함된 발췌문을 챗GPT에 입력했습니다. 이때 챗GPT가 입력값으로 받아들일 수 있도록 문구를 정리하고 가장 불쾌한 텍스트는 제거한 다음 챗GPT에게 불쾌감을 유발하는 문구를 식별하도록 요청해 보았습니다.

챗GPT는 해당 문구를 아주 잘 찾아냈으며 특정 용어의 정의를 언급하면서 해당 용어가 왜 불쾌감을 유발하는지 명확하고 중립적인 방식으로 설명했습니다.

소송 내용이 다소 섬뜩하며 강한 불쾌감을 유발하므로 여기서는 이 내용을 그대로 실지 않기로 했습니다. 원본은 [여기](#)에서 확인할 수 있습니다.

개체를 감지한다

특정 콘텐츠를 넘어, 챗GPT에게 Lionbridge Wikipedia 페이지의 본문에서 전형적인 방식으로 개체를 감지하고 개체 주변에 태그를 달도록 요청해 보았습니다. 두 개를 놓쳤지만 프롬프트 두 개를 더 주었더니 손쉽게 태그를 달았습니다.

나	챗GPT
<p>다음 텍스트에서 관련 용어 주위에 <code><ent type={class}></ent></code> 태그를 추가하세요. 여기서 {class}는 개체 유형입니다.</p> <p>유형은 기업, 브랜드, 위치, 인력, 역할입니다.</p> <p>다음은 텍스트입니다. (...)</p>	<p><code><ent type="company"></code>라이온브리지<code></ent></code>는 1996년에 설립되었습니다. 2005년, 당시 최대 로컬라이제이션 제공업체였던 <code><ent type="company"></code>Bowne Global Solutions<code></ent></code>를 인수했습니다. 2014년, <code><ent type="location"></code>코스타리카<code></ent></code>에 본사를 둔 디지털 마케팅 서비스 에이전시인 <code><ent type="company"></code>Darwin Zone<code></ent></code>과 콘텐츠 연결 소프트웨어 회사인 <code><ent type="company"></code>Clay Tablet Technologies<code></ent></code>를 인수했습니다.</p> <p>2017년 7월, <code><ent type="person"></code>John Fennelly<code></ent></code>가 <code><ent type="role"></code>최고경영자<code></ent></code>로 임명되었습니다. 라이온브리지 창립자인 <code><ent type="person"></code>Rory Cowan<code></ent></code>은 <code><ent type="role"></code>이사회 의장<code></ent></code>을 맡고 있습니다.</p> <p>2020년 11월, <code><ent type="company"></code>라이온브리지<code></ent></code>는 <code><ent type="company"></code> Telus International<code></ent></code>에 AI 사업부를 매각하는 데 합의했습니다.</p>

분류 체계에 따라 사물을 분류한다

챗GPT의 놀라운 특징 중 하나는 특정 상황에 일반적인 지식을 적용할 줄 안다는 것입니다. 챗GPT에게 다음 다섯 문장을 네 가지 범주로 분류할 것을 요청했을 때 챗GPT는 이 일을 원활히 수행하는 한편 왜 그렇게 분류했는지도 설명할 수 있었습니다.



나	챗GPT
<p>문장을 다음 네 가지 범주에 따라 분류하려 합니다.</p> <ul style="list-style-type: none"> • 여행 및 관광 • 석유 및 가스 • 자동차 • 컴퓨터 <p>다음 다섯 개의 문장을 분류해 주세요.</p> <ol style="list-style-type: none"> 1. "TotalEnergies SE, ancienne Compagnie française des pétroles, puis Total, est une compagnie mondiale de production et de fourniture d'énergies" 2. "1936년, 국제연맹은 외국인 관광객을 '최소 24시간 이상 해외를 여행하는 사람'으로 정의했다. 그 후신인 국제연합(UN)은 1945년에 이 정의를 수정하여 체류 기간을 최대 6개월로 지정하는 내용을 포함시켰다." 3. "캐나다의 '퍼스트 네이션(First Nation)' 주민은 선사 시대 초기부터 자작나무껍질로 만든 카누의 방수 처리를 위해 아타바스카 강과 클리어워터 강을 따라 흘러내려온 역청을 사용했다." 4. "이 회사는 1954년 다임러-벤츠 AG와 협력하여 최초의 상용차를 제조했다. 이 차량은 1969년에 생산이 중지되었다." 5. "Apple II는 컬러 그래픽을 표시할 수 있는 기능이 있었으므로 Apple 로고는 다시 디자인되어 무지개 색상을 갖게 되었다." <p>각 문장을 내가 제시한 범주에 따라 분류해 주세요.</p>	<ol style="list-style-type: none"> 1. 석유 및 가스 2. 여행 및 관광 3. 석유 및 가스 4. 자동차 5. 컴퓨터
<p>세 번째 문장을 왜 '석유 및 가스' 범주로 분류했나요?</p>	<p>세 번째 문장을 석유 및 가스로 분류한 이유는 역청을 언급하고 있기 때문입니다. 역청은 도로 포장이나 방수 재료 등에 자주 사용되는 두껍고 검은색이며 끈적끈적한 형태의 석유를 말합니다. 역청은 석유 및 가스 산업의 핵심 구성 요소이므로 문장에 역청이 언급된 것은 이 범주와 관련이 있음을 나타냅니다. 설명이 더 필요한가요?</p>

그렇다면 챗GPT로 무엇을 할 수 있는가?

챗GPT가 이 섹션을 요약한 내용

챗GPT는 번역, 용어집 및 스타일 가이드 준수, 임의의 분류 체계에 따라 텍스트를 분류하고 편집하는 등 언어에 관련된 다양한 작업에서 효과적인 도구입니다. 챗GPT는 로컬라이제이션 업계에서 현재의 기술과 접근법을 도태시키는 파괴력을 보일 수 있습니다. 특히 다양한 언어와 특정 맥락에 대규모로 챗GPT를 도입할 수 있을지 결정하기 위해서는 추가적인 테스트가 필요합니다.

첫 번째 섹션에서는 챗GPT에 의존해서는 그 답이 참인지 무엇이 옳은지 알 수 없다는 것을 확인했습니다. 다시 말해, 이 기술을 사용하려는 콘텐츠 작성자는 챗GPT가 터무니없는 답변을 한 것은 아닌지 검토해야 한다는 어려움이 있습니다. 무엇이 말이 되는지, 무엇이 참인지 거짓인지, 어떤 의견을 표하고 싶은지 결정하는 것은 여전히 인간의 몫입니다. 세상을 이해하는 일은 인간 혼자서 감당해야만 하는 짐입니다.

그러나 만족스러운 의미가 포함된 텍스트를 받으면 챗GPT는 암호화된 의미는 유지한 채 텍스트의 형식과 내용을 조작하거나 변환할 수 있습니다.

이는 로컬라이제이션 업계 종사자에게는 좋은 기회입니다. 로컬라이제이션에 필요한 의미 있는 콘텐츠를 처음부터 새로 만들 필요가 없기 때문입니다. 로컬라이제이션의 목표는 기존 콘텐츠를 분석하고 변환하며 개선하는 것입니다. 이제 로컬라이제이션 작업 환경과 챗GPT가 오늘날 로컬라이제이션 업계에 어떤 영향을 미칠 수 있는지 자세히 살펴보겠습니다.

챗GPT는 번역에 유용하다

무엇보다도 콘텐츠 번역에 활용도가 높습니다. 챗GPT는 번역에 매우 뛰어납니다. 시간이 지나면 다양한 언어로 번역 범위가 확대될 것으로 예상됩니다. 물론 언어마다 격차와 고유한 특징은 있겠지만, 말뭉치가 큰 언어의 경우 챗GPT는 비록 우수하지 않은 점도 일부 있지만 뛰어난 품질의 기계번역 결과를 보여주는 최첨단 엔진과 맞먹는 수준에 이를 수 있습니다.

챗GPT는 용어 관리에 유용하다

둘째, 번역을 할 때는 다른 용어보다 특정 용어를 더 우선적으로 사용해야 하는 경우가 많습니다. 이를 위해 번역 제공업체는 용어집을 비롯한 다양한 용어 관리 기법을 활용합니다. 기계번역의 경우, 용어 관리 문제는 아직 완전히 해결되지 않고 있습니다. 여기에는 다양한 접근방식이 있으며 상황에 따라 다르게 적용되지만 다음과 같이 요약할 수 있습니다.

- 핵심적인 기계번역 워크플로에 부가 기능으로 용어집을 삽입하거나
- 특별히 선별된 콘텐츠 세트와 용어집을 사용하도록 엔진을 학습시킴

이러한 기법을 단독으로 사용하거나 함께 사용하면 효과가 있지만 용어집이 올바르게 적용되었는지의 문제는 번역 품질 관리에서 여전히 핵심 사안으로 남아 있습니다. 이 두 가지 접근방식은 용어 품질에 빈틈을 만들 수 있고 이러한 기술과 말뭉치를 관리하는 데는 엄청난 노력이 소모되므로 이 기법을 사용하는 데 있어 경제적 효과는 제한적입니다. 반면 챗GPT는 앞의 몇 가지 사례를 통해 볼 수 있었던 것처럼 상당히 효과적으로 용어 지침을 준수하는 것으로 보입니다.

챗GPT는 스타일가이드에 유용하다

셋째, 스타일 지침을 넓게(예: 비격식체로 작성) 또는 좁게(예: 독자를 지칭할 때 2인칭 복수형 사용) 적용할 수도 있습니다. 반면, 기계번역 엔진은 이러한 종류의 작업을 제대로 처리하기 어렵습니다. 챗GPT는 이 작업에 꽤 능숙한 것으로 보이며 적용하기도 간단합니다.(나중에 자세히 설명하겠습니다.)



챗GPT는 콘텐츠 분류에 유용하다

넷째, 분류 작업에 뛰어나며 특히 텍스트에 임의 분류 방식을 적용하는 데 능숙합니다. 특정 유형의 콘텐츠별로 각기 다른 지침을 적용할 수 있으므로 로컬라이제이션에 특히 유용합니다. 예를 들어 콘텐츠가 X에 관한 내용이면 용어집 Y를 적용하고, 문서라면 격식체를 사용하는 식으로 설정할 수 있습니다. 반면, 기존의 번역 메모리(TM)는 주제별 적용이 상당히 어려운 것으로 알려져 있으며, 어떤 지침을 적용해야 할지 결정을 내려야 하는 상황에는 사람의 판단에 의존할 수밖에 없습니다. 이러한 결정을 자동화할 수 있는 신뢰할 수 있는 방법이 있다면 매우 유용할 것입니다.

챗GPT는 사후 편집에 유용하다

다섯째, 챗GPT의 놀라운 장점 중 하나는 주어진 텍스트를 편집하는 능력입니다. 텍스트 편집, 구체적으로 사후 편집은 로컬라이제이션 품질을 개선하는 데 가장 중요한 요소입니다. 편집은 번역된 텍스트를 검토하고 다음 네 가지를 확인하는 작업으로 구성됩니다.

1. 번역된 텍스트가 원문과 동일한 의미를 전달하는가
2. 올바른 언어가 사용되었는가
3. 스타일과 어조가 적절한가
4. 필수 용어가 사용되었는가

챗GPT는 이 모든 작업에 상당히 능숙한 것으로 보입니다.

챗GPT는 콘텐츠 분석에 유용하다

마지막으로, 효과적인 처리, 개선, ROI를 위한 텍스트 분석을 지원합니다. 텍스트 분석의 목표는 번역 품질 문제를 예측 또는 사전 파악하고, 도달범위, 검색 엔진 최적화(SEO), CTA/CTR 실적 등 결과의 효율성에 맞춰 조정하며, 원문 및 대상 콘텐츠의 가독성을 개선하는 것입니다. 챗GPT는 텍스트를 분석하고 그 결과를 명확하게 보고할 수 있으므로 이러한 기능을 콘텐츠를 필터링하거나 번역된 콘텐츠를 개선하는 데도 사용할 수 있습니다.



챗GPT는 작업 코드 작성 및 편집에 유용하다

챗GPT에는 아직 언급하지 않은 또 다른 핵심 기능이 있습니다. 바로 코드를 작성하고 편집하는 기능입니다. 원하는 작업을 명확하게 기술한 후 코드를 작성할 것을 요청하면 챗GPT는 그러한 요청에 따라 코드를 작성합니다. 최정예 프로그래머들이 챗GPT가 자신들의 수준만큼 코드를 작성할 수 있을지를 두고 논쟁을 벌였는데 결과는 **그렇지 않아 보입니다**. 그러나 다음과 같은 프롬프트를 입력했을 때 챗GPT는 몇 가지 콘텐츠 추출 코드를 작성할 수 있었습니다.

나(여러 프롬프트)

로컬 파일 시스템의 xml 파일에서 콘텐츠를 추출하는 python 스크립트를 작성해 주세요. 추출하려는 콘텐츠는 다음 XML 태그에 포함시킵니다. 추가적인 특수 지침은 문맥 중 {이런 방식} 으로 중괄호를 사용하여 표시하겠습니다. 다음은 콘텐츠를 추출해야 하는 태그의 정의입니다.

```
<maincontent>{다른 태그 내 모든 콘텐츠 추출} </maincontent>
```

추출된 콘텐츠는 'content.txt'라는 이름의 파일로 출력합니다.

작성된 스크립트를 이용하면 메인 콘텐츠 태그의 첫 번째 수준에 있는 콘텐츠만 추출됩니다. 하위 태그에서도 콘텐츠를 추출해야 합니다.

작성한 스크립트를 사용하면 메인 콘텐츠 요소 내용 중 두 번째 수준 태그 이전의 첫 번째 부분만 추출됩니다. 두 번째 수준 태그 이후의 두 번째 부분은 추출되지 않습니다.

이제 작동합니다. 하지만 콘텐츠를 작성한 후 파일을 정리할 필요가 없었으면 좋겠습니다. 그냥 처음부터 빈 줄 없이 작성해 주세요.

챗GPT(최종 결과물)

```
import xml.etree.ElementTree as ET
# XML 파일 구문 분석
tree = ET.parse('file.xml')
# 루트 요소 가져오기
root = tree.getroot()
# 쓰기 모드에서 새 파일 열기
with open('content.txt', 'w') as f:
    # 모든 'maincontent' 요소 찾기
    for maincontent in root.findall('maincontent'):
        # 'maincontent' 요소의 모든 하위 요소 반복
        for child in maincontent.iter():
            # 하위 요소 내부의 콘텐츠 추출
            content = child.text
            # 콘텐츠가 비어 있지 않은지 또는 공백 문자로만 구성되어
            # 있는지 확인
            if content.strip():
                # 파일에 콘텐츠 작성
                f.write(content)
                # 콘텐츠 작성 후 줄바꿈 문자 추가
                f.write("\n")
            # 하위 요소에 'tail' 속성이 있는지 확인
            if child.tail is not None:
                # 'tail' 속성 추출
                tail = child.tail
                # 'tail' 속성이 비어 있지 않은지 또는 공백 문자로만
                # 구성되어 있는지 확인
                if tail.strip():
                    # 파일에 'tail' 속성 작성
                    f.write(tail)
                    # 'tail' 속성 작성 후 줄바꿈 문자 추가
                    f.write("\n")
```

이 스크립트는 요청한 바에 따라 잘 작동했습니다. 코드를 복사하여 .py 파일에 넣고 샘플 콘텐츠가 포함된 XML 모의 파일을 만든 다음 이를 실행하면 올바른 콘텐츠가 추출되어 content.txt 파일로 만들었습니다. 추가 프롬프트가 표시될 때마다 이 파일을 실행하고 챗GPT에게 무엇이 잘못되었는지 알려주면 챗GPT는 스크립트를

업데이트했습니다. 아마도 더 적은 프롬프트 줄로 동일한 작업을 수행할 수 있는 방법이 있으며 파일이 복잡해지면 버그도 발생할 것이라고 생각합니다. 그러나 프로그래머가 아닌 사람도 코드 한 줄 입력하지 않고도 몇 분 만에 이 작업을 간단히 수행할 수 있었습니다.

다음과 같이 XML 파일을 입력했을 때 아래의 content.txt 파일이 출력되었습니다.

나	챗GPT
<pre><xml> <headerstuff>이 제목은 추출하고 싶지 않음</headerstuff> <maincontent>내가 원하는 콘텐츠입니다. <secondlevel>두 번째 수준의 콘텐츠입니다. </secondlevel> 이 또한 두 번째 수준 태그 다음에 원하는 콘텐츠입니다. <secondlevel>또 다른 두 번째 수준의 콘텐츠입니다.</secondlevel> 마지막 tail입니다. </maincontent> <othercontent>원하지 않는 태그의 콘텐츠입니다. </othercontent> <othercontent>상위 요소에 표시되는 것을 원하지 않는 콘텐츠입니다. </xml></pre>	<p>내가 원하는 콘텐츠입니다. 두 번째 수준의 콘텐츠입니다.</p> <p>이 또한 두 번째 수준 태그 다음에 원하는 콘텐츠입니다. 또 다른 두 번째 수준의 콘텐츠입니다.</p> <p>마지막 tail입니다.</p>

챗GPT는 스크립트가 무슨 작업을 수행하는지 명확하게 설명하는 주석을 코드에 표시했으며, 요청하면 자세히 설명할 수 있습니다(여기에는 표시하지 않음). 컴퓨터에서 코드를 실행하는 방법 등에 대해서도 조언을 얻을 수 있습니다. 이를 통해 코드 작성에 문외한인 사람도 쉽게 코드 작성법을 배우고 코드를 만들 수 있습니다.

제작에 활용하여 기대하는 결과를 얻으려면 효과적으로 프롬프트할 수 있는 전문 기술을 개발해야 합니다. 콘텐츠의 특정 변환에는 정리, 사전 처리 및 사후 처리 등 각각 다른 작업을 수행하는 일련의 프롬프트가 필요할 수 있습니다.

자동화 파이프라인 구축의 일환으로 문맥 관련성이 높고 결과물에서 충분히 예측할 수 있는 방식으로 자연어 프롬프트 사용법을 확립하는 것도 흥미로운 과정이 될 것입니다. 이제는 학습 과정뿐만 아니라 제작 과정의 자동화 감독도 더 이상 결과물을 평가하고 주석을 다는 것에 국한되지 않으며, 입력 과정을 감독해야 하는 경우도 생길 수 있습니다.

챗GPT가 규정할 새로운 기술 및 관행

지금까지 살펴보았듯이, 챗GPT와 다양한 주제로 길고 복잡한 대화를 나누어 보니 이제 원하는 결과를 얻기 위해서는 챗GPT에 어떤 유형의 프롬프트가 필요한지 이해할 수 있습니다. 챗GPT는 낯선 언어에서 모국어에 이르기까지 모든 형식의 언어로 지침을 이해한다는 점에서 깊은 인상을 받았지만, 원하는 작업을 수행하기 위해 프롬프트를 여러 번 다시 작성할 필요가 있었습니다. 그중 몇 가지는 여기서 자세히 보여주었으나, 프롬프트를 여러 번 재구성한 이후에야 이 백서에 예제로 실을 정도인 경우도 있었습니다. 이런 과정을 통해 우리는 챗GPT와 어떻게 대화해야 하는지 이해할 수 있었습니다. 이 특별한 기술은 자연어만 입력값으로 받아들입니다. 따라서 이 기술을

일련의 일반 프롬프트를 광범위한 콘텐츠 및 작업 영역에 적용할 수 있을까요? 프롬프트에 대한 자연어 문맥을 자동으로 정의하고 예상되는 작업을 수행하는 동안 문맥에 맞춰 조정하려면 어떻게 해야 할까요? 이제 새로운 프로젝트마다 QA 및 사후 편집 중 챗GPT 프롬프트를 조정하기 위한 새로운 엔지니어링 작업이나 언어 작업을 수행해야 하나요? 이러한 질문에는 여전히 대답하기 어렵습니다.



향후 로컬라이즈 업계가 나아가야 할 방향

챗GPT는 보통의 인간 편집자와 대등한 수준으로 텍스트를 변환하고 주석을 달 수 있으며 이러한 작업을 더 효율적으로 처리할 수 있습니다. 사실상 사람에게 없는 다양한 기술을 활용하여 작업할 수 있으며, 보유한 지식을 가지고 새로운 상황에 일반화할 수도 있습니다.

가장 중요한 점은 자동 번역에 특정 용어집 및 스타일 요건을 적용하는 것과 같이 일부 남아 있는 로컬라이제이션 자동화 문제도 챗GPT가 해결할 조짐이 보인다는 사실입니다.

물론 소규모 연습용 데이터를 사용해 챗GPT와 대화를 나누고 원하는 작업을 수행하도록 하는 것도 한 방법입니다. 그러나 이러한 작업에 챗GPT를 대대적으로 사용하려 한다면 그것은 별개의 사안입니다. 늘 그렇듯이, 아이디어 도출이나 개념 검증(POC)은 크게 어렵지 않습니다. 이 기술을 대규모 가치를 창출할 수 있도록 전환하는 것이 지금 해결해야 할 과제입니다.

미래로 나아가기 위해서는 다음과 같은 과제를 처리해야 합니다.

- 대규모로 실제 테스트를 실행하여 여기에서 조사한 각 유형의 로컬라이제이션 및 편집 작업의 오류율을 평가한다.
- 로컬라이제이션 가치 사슬 내에서 발생하는 매크로 및 마이크로 사용자 여정을 면밀히 분석하고 이러한 종류의 텍스트 자동화로 어디에 문제가 생길 수 있는지를 파악한다.

- 챗GPT에 프롬프트를 입력하고 관련 컨텍스트를 대량으로 제공하는 방법을 이해하고, 위험 요소와 모범 사례를 문서화한다.
- 새로운 자동화 프로세스 및 인간 참여형 편집 워크플로를 개발하고, 앞으로 이러한 AI 참여형 작업이 부상할 미래의 사후 편집과 QA가 의미하는 바를 간파한다.
- 개선 가능성이 있는 모든 기회를 대상으로 로컬라이제이션 업체 및 고객 모두를 위한 새로운 자동화 및 사용자 경험(UX) 상호작용 환경을 설계한다.
- 비즈니스의 규모에 맞춰 라이선스, 배포 및 유지관리 비용을 효과적으로 운용한다.

새로 등장한 이 생성형 AI가 무시할 수 없는 존재라는 것은 분명합니다. 챗GPT는 로컬라이제이션 업계를 큰 혼란에 빠뜨릴 수 있습니다. 챗GPT의 실질적인 잠재력을 파악하고 비즈니스에서 가치를 창출하기 위해서는 위의 과정들을 수행하는 것이 필수입니다.

이러한 기술이 점차 보편화됨에 따라 이 기술을 사용하는 고객과 공급업체도 늘어나게 될 것입니다. 따라서 뒤처지지 않기 위해서는 언어 자동화를 주도적으로 추진해야 합니다.

챗GPT의 언어 및 실제 사용에 대한 고려 사항

챗GPT를 사용해 보고 느낀 가장 놀라운 점은 복소수 연산을 거의 정확하게 처리하면서도 여전히 연산 결과는 틀린다는 것입니다. 이는 챗GPT가 속임수를 쓰지 않는다는 점을 보여줍니다. 실제로 챗GPT는 훈련된 언어를 통해 모든 것을 학습합니다. 무언가를 계산해야 한다는 것을 감지했을 때도 계산기를 두드리지 않습니다. 수에 대한 개념은 있지만 텍스트로부터 수 개념을 학습했으므로 매우 기초적인 연산 이외에는 정확한 결과를 내지 못합니다.

챗GPT가 특정 자릿수를 넘어가는 대규모의 작업에서 거의 정확한 결과값(숫자가 작다면 정확한 결과값)을 찾아낼 수 있다는 사실은 규모가 충분히 큰 언어 말뭉치의 경우 실제 세계에 대한 통계적으로 의미 있는 지식이 들어 있음을 시사합니다.

챗GPT는 연산 결과를 추정하기에는 충분합니다. 그러나 실제 세계에 대한 유의미하고 신뢰할 수 있으며 정확한 정보를 생성하기 위해서는 **전문 공식 체계(예: 수학 공식)** 역시 필요하다는 사실도 알 수 있습니다.

이것은 챗GPT가 자신의 응답을 적절히 정당화할 수 있지만 이를 제대로 적용하지 못한다는 점에서도 확인할 수 있습니다.

챗GPT가 작동하는 방식에서 언어 그 자체와, 지식을 생산하는 데 언어가 하는 역할에 대해 몇 가지 통찰을 얻을 수 있습니다.

- **언어는 그 자체로 인간이 세상을 어떻게 이해하는지에 관한 대부분의 사실을 담고 있습니다.** 언어의 자기 일관성은 언어가 반(半) 형식적인 의미 생성 기계에 불과함을 보여줍니다. 이러한 현실로 인해 챗GPT는 콘텐츠가 주어졌을 때 동일한 의미 수준을 유지하면서 원하는 만큼 내용을 수정할 수는 있지만 새로운 진짜 지식을 안정적으로 생성할 수는 없습니다. 이러한 상황은 어떤 속설을 상기시킵니다. 즉, 자기 일관적인 패턴은 지식과 비슷한 무언가를 만들 수는 있지만 현실의 통제에서는 벗어날 수 없습니다.

- **반대로, 진리를 확립하려면 언어는 언어의 구문론 및 의미론과는 독립적으로 존재하는 실제 객체와 정의된 기호를 연산하는 형식적인 하위 시스템에 뿌리를 내려야 합니다.** 무엇이 진리인지에 대한 궁극적인 심판자는 현실 그 자체이며, 언어는 의미 조작과 의사소통의 껍질로 현실을 둘러싸고 있습니다. 적절한 언어를 논리적으로 사용하면 터무니없는 소리조차 겉보기에 자기 일관적인 것으로 들리도록 할 수 있으므로 매우 설득력 있게 만들 수 있습니다.
- **챗GPT는 진리는 현실 세계의 속성인 반면, 의미는 우리가 머릿속에서 만들어낸 것임을 일깨워줍니다.** 의미는 언어 그 자체가 아니며, 우리의 광범위한 인지 및 감각 장치의 한 속성일 뿐입니다.

지금 우리는 우리의 마음이 언어와 문화를 통해 세상을 만들었다는 이론을 자주 접하곤 합니다. 동시에, 이러한 주장의 반대편에서는 우리가 신경망과 언어 모델을 사용하여 세상을 이해하는 일반화된 인공지능과 의식이 있는 기계를 만들 수 있다고 말하는 사람들이 있습니다.

챗GPT는 스스로 참조하고 일관성을 갖춰가는 시스템이 그 자체만으로는 독립적으로 존재하는 세상의 진실을 전달할 수 없다는 사실을 냉철하게 일깨워줍니다. 마치 '참이지만 증명할 수 없는 명제가 있다'는 괴델의 불완전성 정리를 떠올리게 합니다. 인간에게 주어진 실제 세계는 언어만으로는 정의할 수 없으며 형식적, 물질적인 시스템은 토대로 이해할 수 있으므로 의식이 있는 존재로서의 우리는 우리의 인식을 이러한 시스템으로부터 분리할 수 없습니다. 챗GPT는 이 수수께끼를 풀지 못하며, 특이점은 여전히 우리 손에 닿지 않는 곳에 있습니다.



라이온브리지 소개

라이온브리지는 장벽을 허물어 세상을 연결하기 위해 다양한 브랜드와 파트너십을 맺고 있습니다. 라이온브리지는 25년 넘게 350개 이상의 언어로 번역 및 로컬라이제이션 솔루션을 제공하여 고객사가 자사 직원 및 글로벌 고객과의 연결성을 높일 수 있도록 도왔습니다. 라이온브리지는 세계적인 수준의 플랫폼을 기반으로 전 세계의 열정적인 전문가 네트워크를 구축하여 다양한 브랜드와의 파트너십을 통해 문화적으로 풍부한 경험을 제공할 수 있습니다. 또한 언어학에 대한 꾸준한 열정을 바탕으로 고객의 고객, 즉 최종 사용자와 공감대를 형성할 수 있도록 최상의 인간 지능과 기계 지능을 두루 활용합니다. 본사는 미국 매사추세츠주 월섬에 있으며 23개국에서 솔루션 센터를 운영하고 있습니다.



자세한 내용

LIONBRIDGE.COM

LIONBRIDGE

© 2023 Lionbridge. All Rights Reserved.