

LESSON

초급 초6·중2 AI 교실

08

AI도 틀릴 수 있어요

AI는 짐작해서 답해요. 그래서 틀릴 수 있고, 중요한 일은 꼭 확인해요.

 **학습 목표** 이번 시간이 끝나면...

- ✓ AI가 틀리는 4가지 모습을 예로 들 수 있어요.
- ✓ AI가 자신 있게 말해도 틀릴 수 있는 까닭을 설명할 수 있어요.
- ✓ AI의 답을 믿을 만한 자료로 확인하는 방법을 말할 수 있어요.

• 초급 08

AI도 틀릴 수 있어요

AI는 짐작해서 답하기 때문에 틀릴 수 있어요.

중요한 일은 꼭 한 번 더 확인해요.



“비 온다더니
해가 짹짹하네?”

AI도 짐작이라서 빗나갈 수 있어요.

⚠ AI가 틀리는 4가지 모습

1



사진
착각

2



번역
실수

3



지어낸
답

4



잘못
알아듣기

📖 꼭 알아야 할 단어

환각

= AI가 없는 사실을 그럴듯하게 지어내는 것

사실 확인

= 맞는 정보인지 믿을 만한 자료로 확인하기

Q. AI가 자신 있게 말하는 답은
항상 맞는다?

O

/

X

정답: X

그림 1 한눈에 보는 08강 — AI가 틀리는 네 가지 모습을 먼저 살펴보세요.



쉽게 풀어 보기

AI가 왜 틀리는지, 틀릴 때 어떻게 하면 되는지 알아봐요.



AI는 왜 틀릴까요?

AI는 정답을 ‘알고’ 말하는 게 아니라, 배운 것을 바탕으로 **가장 그럴듯한 답을 짐작**해요. 짐작은 맞을 때도, 틀릴 때도 있어요.

특히 조심할 점이 있어요. AI가 없는 사실을 그럴듯하게 지어내는 것을 **환각(hallucination)**이라고 해요. 없는 책 제목이나 틀린 날짜를 진짜처럼 말하기도 해요.

그래서 맞는 정보인지 믿을 만한 자료로 확인하는 **사실 확인(fact check)**이 꼭 필요해요.



빗나간 일기 예보로 이해해요

“내일 비 온대” 하고 우산을 챙겼는데 해가 뽀뽀했던 적 있나요? 일기 예보는 지금까지 모은 날씨 자료로 내일 날씨를 짐작하는 거라 가끔 빗나가요.

AI도 마찬가지예요. 많이 배워서 대부분 잘 맞히지만, **짐작이기 때문에 언제든지 틀릴 수 있어요.**

더 조심할 점은 AI가 **틀린 답도 아주 자신 있는 말투로** 말할 때가 있다는 거예요.



한 줄 요약

AI는 배운 것을 바탕으로 짐작해서 답하기 때문에 틀릴 수 있어요. 그래서 중요한 일은 꼭 한 번 더 확인해야 해요.

📖 생활 속에서 AI가 틀릴 때

번역기가 우리말 관용 표현을 글자 그대로 옮기거나, AI 비서가 시끄러운 곳에서 말을 잘못 알아듣는 일은 흔해요.

챗봇에게 숙제를 물었을 때 그럴듯하지만 틀린 설명을 내놓기도 해요. 말투가 자연스럽다고 내용까지 맞는 것은 아니에요.

확인하는 방법은 어렵지 않아요. **책이나 믿을 만한 누리집에서 찾아보기, 선생님이나 보호자에게 여쭙보기**예요. 건강이나 안전처럼 중요한 일일수록 꼭 확인해요.

미리 보기 뒤에 나오는 '직접 해보기 미션'에서 번역기가 관용 표현을 어떻게 옮기는지 직접 확인해 봐요.

미리 생각하기 AI나 앱이 틀린 답을 낸 것을 본 적이 있나요? 있었다면 적어 보세요.

예: 길 안내가 이상했어요, 내 말을 잘못 알아들었어요.

123 숫자로 보는 08강

예시

4가지

이 강에서 살펴본 AI가 틀리는 모습: 사진 착각, 번역 실수, 지어낸 답, 잘못 알아듣기

예시

10번 중 9번

대부분 맞아도 한 번은 틀릴 수 있어요. 예로 든 숫자예요.

예시

2곳 이상

중요한 정보는 믿을 만한 자료 두 곳 이상에서 확인하면 좋아요.

사실 = 실제로 있었던 일 · **예시** = 이해를 도우려고 만든 숫자



표로 정리해요

AI가 틀리는 모습과 확인하는 방법을 정리해요.



AI가 틀리는 4가지 모습

모습	이렇게 틀려요	이렇게 확인해요
① 사진 착각	초코칩 머핀 사진을 강아지 얼굴로 잘못 알아봐요.	내 눈으로 사진을 다시 봐요.
② 번역 실수	“발이 넓다(아는 사람이 많다)”를 “발 크기가 크다”는 뜻으로 옮겨요.	거꾸로 번역해 뜻을 비교해요.
③ 지어낸 답	없는 책 제목이나 틀린 날짜를 진짜처럼 말해요.	책, 믿을 만한 누리집에서 찾아봐요.
④ 잘못 알아듣기	시끄러운 곳에서 내 말을 다른 말로 알아들어요.	조용한 곳에서 또박또박 다시 말해요.



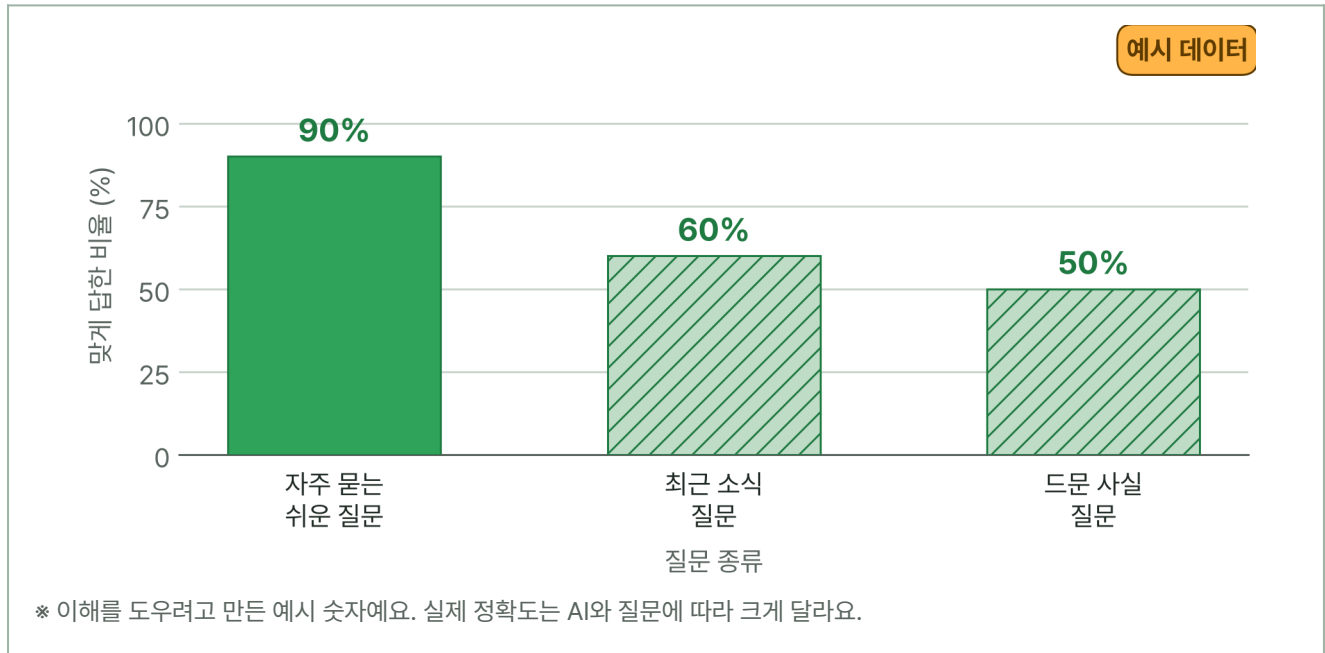
믿을 만한 확인과 믿기 어려운 확인

살펴봐요	믿을 만한 확인	믿기 어려운 확인
어디서	책, 공공 기관 누리집, 백과사전	출처를 모르는 글, 소문
누구에게	선생님, 보호자, 전문가	확인하지 않은 친구의 말
어떻게	두 곳 이상에서 비교해요.	마음에 드는 답이 나올 때까지 다시 물어요.

생각해 보기 AI가 알려 준 역사 날짜를 확인하려면 어디에서 찾아볼 수 있을까요? 두 곳을 적어 보세요.

교과서, 도서관 책, 믿을 만한 백과사전 누리집을 떠올려 봐요.

질문 종류에 따라 다른 AI 답의 정확도 (예시)



그래프 읽기 어떤 질문에서 AI가 가장 많이 틀렸나요? 그런 질문일수록 어떻게 해야 할까요?

이 그래프는 예시 데이터예요. 막대가 낮을수록 확인이 더 필요해요.

가 꼭 알아야 할 단어

환각

hallucination

AI가 없는 사실을 그럴듯하게 지어내는 것

사실 확인

fact check

맞는 정보인지 믿을 만한 자료로 확인하기

**직접 해보기 미션**

손과 눈으로 직접 확인해 봐요.

번역기로 관용 표현 확인하기

가입하지 않아도 쓸 수 있는 번역기 누리집(예: 파파고, 구글 번역)에 **“그 친구는 발이 넓어.”**를 넣어 영어로 바꿔 보세요. 나온 영어 문장을 다시 한국어로 번역해서 처음 뜻과 같은지 비교해 봐요.

- 1 “그 친구는 발이 넓어.”를 영어로 번역해요.
- 2 나온 영어 문장을 다시 한국어로 번역해요.
- 3 번역이 잘 나왔다면 “손이 크다”처럼 다른 표현도 넣어 보세요.

※ 결과는 번역기마다, 때마다 다를 수 있어요. 이름이나 주소 같은 개인 정보는 입력하지 않아요.

기록표

	영어로 번역한 결과	다시 한국어로	뜻이 같나요?
발이 넓다			
손이 크다			
내가 고른 표현			

- 뜻이 달라진 표현이 있었나요? 왜 그렇게 되었을까요?

잘 번역된 표현**이상하게 번역된 표현**

🔑 정리 AI 번역도 틀릴 수 있어요. 중요한 글은 사람이 꼭 확인해요.

**확인 퀴즈**

스스로 풀고, 정답은 마지막 쪽에서 확인해요.

1. **O/X** AI가 자신 있게 말하는 답은 항상 맞는다.☐

O

☐

X

2. **고르기** AI의 답을 확인하는 방법으로 알맞지 않은 것은?☐

① 책이나 믿을 만한 누리집에서 찾아보기

☐

② 선생님이나 보호자에게 여쭙보기

☐

③ 마음에 드는 답이 나올 때까지 다시 물어보고 그 답을 믿기

3. **빈칸** AI가 없는 사실을 그럴듯하게 지어내는 것을 _____(이)라고 해요.4. **쓰기** 건강에 관한 정보를 AI에게 들었다면 어떻게 해야 할까요? 한 문장으로 써 보세요.

**중2 한 걸음 더**

생성형 AI는 왜 그럴듯하게 틀릴까?

챗봇 같은 **생성형 AI(generative AI)**는 문장에서 다음에 올 말로 가장 그럴듯한 것을 확률로 골라 이어 가요. 그래서 문장은 자연스러워도 내용은 틀릴 수 있어요.

또 학습한 시점 이후의 소식은 모를 수 있고, 배운 데이터가 치우쳐 있으면 답도 치우쳐요. 속제나 건강, 안전처럼 중요한 일일수록 **출처를 확인하고, 마지막 판단은 사람이** 해야 해요.

- 생각해 봐요: AI의 답에 출처(어디서 나온 정보인지)가 함께 있으면 왜 도움이 될까요?



용어 정리

이번 시간에 나온 어려운 말을 모았어요.

용어	영어	쉽게 말하면	예시
환각	hallucination	AI가 없는 사실을 그럴듯하게 지어내기	없는 책 제목 말하기
사실 확인	fact check	맞는 정보인지 믿을 만한 자료로 확인하기	백과사전에서 날짜 찾기
짐작	guess	확실하지 않지만 그럴 것이라고 생각하기	"내일 비가 올 거야."
출처	source	정보가 나온 곳	책 제목, 누리집 주소
관용 표현	idiom	글자 뜻과 다른 뜻으로 쓰는 말	"발이 넓다"
번역	translation	한 나라 말을 다른 나라 말로 바꾸기	한국어 → 영어
음성 인식	speech recognition	목소리를 글자로 바꾸는 기술	AI 비서가 내 말 알아듣기
생성형 AI	generative AI	글, 그림 같은 새 결과물을 만들어 내는 AI	챗봇
확률	probability	어떤 일이 일어날 것 같은 정도	"비 올 확률 60%"
누리집	website	인터넷에서 볼 수 있는 홈페이지	번역기 누리집



정답과 풀이

퀴즈를 다 푼 다음에 확인해요

1. X — AI는 틀린 내용도 자신 있는 말투로 말할 수 있어요.
2. ㉠ — 마음에 드는 답이 곧 맞는 답은 아니에요.
3. 환각 — 없는 책 제목을 진짜처럼 말하는 것도 환각이에요.
4. (예시 답) 그대로 믿지 말고 보호자나 의사 선생님께 여쭙거나 믿을 만한 자료로 확인해요.