



경기과학기술대학교 · 교양(AIDX)

시프롬프트 디자인

7주차

그럴듯한 답일수록 밖에서 직접 확인한다

글이 자연스럽게 보여도
모두 맞는 말이 아닐 수 있다.

오늘의 강의

1교시에 왜 틀리는지 보고, 2교시에 밖에서 확인하는 절차와 검증 실습

1교시



자신 있게 틀린다

환각의 원인과 모양
묻고 바로 확인하기

2교시



직접 확인하는 절차

출처 · 원문 · 다시 세기
내 기획안 검증 표

1교시

자신 있게 틀린다

그럴듯한 답이 왜 틀리는지 보고, 묻자마자 확인해 본다.

환각

원인 셋과 모양 셋

생활 사례 넷

지난주는 계획 안에서, 오늘은 밖에서 찾는다

6주차 점검

계획 안에서 빠진 것, 부딪치는 것, 조건을 넘는 것

7주차 검증

받은 사실이 맞는지 밖에서 직접 찾아 확인

지난주 만든 네 단계 기획안을 열어 두자.

우리 학교 처음 이름과 지금 이름이 된 해를 물으면

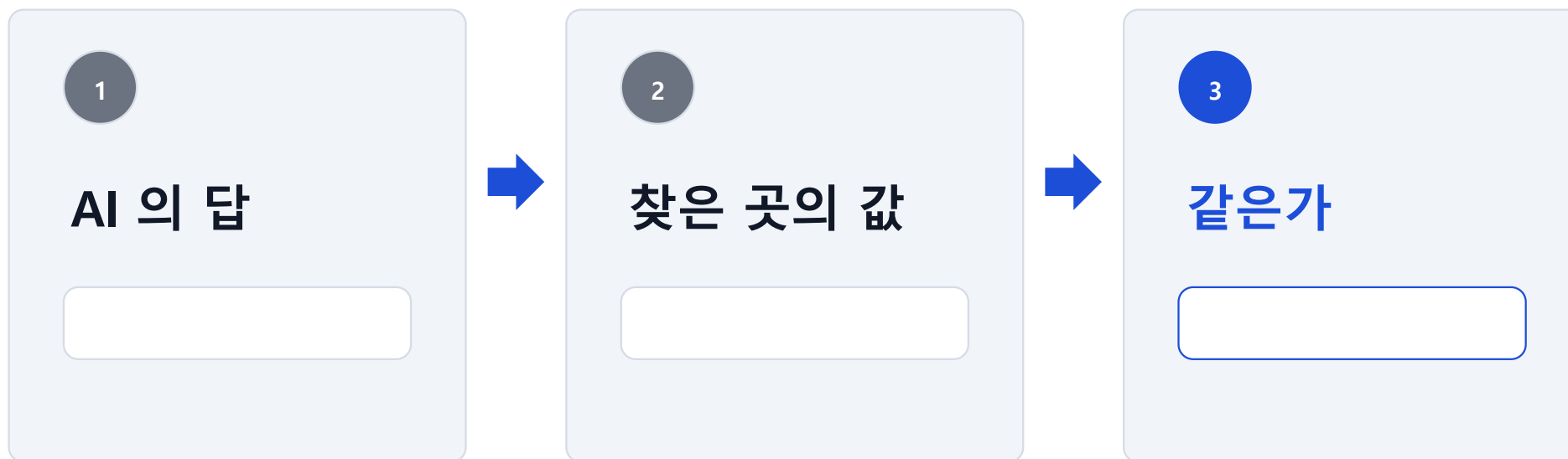
경기과학기술대학교가 처음 문을 열 때 학교 이름은 뭐였어?
지금 이름인 「경기과학기술대학교」가 된 것은 몇 년이야?

ChatGPT · GPT-5.6 Sol

여기서 실제 화면으로 함께 본다

학교 연혁에서 직접 찾으면

확인할 곳: 경기과학기술대학교 · 대학안내 · 연혁



환각

Hallucination

환각(Hallucination): AI 가 사실이 아닌 정보를 자신감 있게 제시하는 현상.

사실과 다른 내용을 사실처럼 확신에 찬 어조로 생성하는 현상을 할루시네이션(Hallucination, 환각)이라 하며, 원인은 초능력의 고장이 아니라 <다음 토큰 예측>에서 배운 작동 원리 그 자체에 있다.

거대한 앵무새, 사실 확인에 약하다

AI의 한계 (앵무새)

사실 확인에 취약함 (환각 현상)

맥락 없이는 의도 파악 불가

AI의 가능성 (창의적 파트너)

방대한 데이터를 활용한 지식의 확장 및
연결

명확한 맥락 제공 시, 정확하고 깊이 있는
결과

AI는 신념이나 감정, 자의식이 없으며, 때로는 학습한 데이터에 없는 내용을 사실처럼 지어내는 '환각 (Hallucination)' 현상을 보이기도 한다.

왜 틀리나, 원인 셋

확률로 고른다

확률적 생성, 사실 조회 아님: LLM은 데이터베이스를 뒤져 답을 찾는 것이 아니라 "이 문맥에 이어질 확률이 높은 토큰"을 뽑는다

모르는 줄 모른다

지식의 경계가 없음: 신입사원은 자기가 우리 회사 규정집을 읽은 적 없다는 사실을 "알지" 못한다.

시간이 멈췄다

시간이 멈춰 있음: 모델의 지식은 학습 데이터가 수집된 시점에서 멈춘다.

「경조사 휴가는 ○일입니다」 틀에 숫자를 채운다

경조사 휴가는 ?일입니다

“경조사 휴가는 며칠인가?”라는 질문 뒤에 이어질 법한 문장의 모양(예: “경조사 휴가는 ○일입니다”)은 학습 텍스트에 무수히 있었으므로, 모델은 그 틀에 그럴듯한 숫자를 채워 넣는다.



시험에서 빈칸보다 찍기가 점수를 더 받는다면

If you do not know the answer but take a wild guess, you might get lucky and be right. Leaving it blank guarantees a zero. (정답을 모르더라도 무작정 찍어보면 운 좋게 맞힐 수도 있습니다. 하지만 빈칸으로 남겨두면 확실히 0점이 됩니다.)

Over thousands of test questions, the guessing model ends up looking better on scoreboards than a careful model that admits uncertainty. (수천 개의 시험 문항을 놓고 보면, 불확실성을 인정하는 신중한 모델보다 추측에 의존하는 모델이 결국 점수판에서 더 좋은 성적을 거두는 것으로 나타납니다.)

모르는 문제를 찍으면 맞을 수도 있다. 빈칸으로 두면 0점이다.

수천 문제를 보면, 찍는 모델이 모른다고 말하는 모델보다 점수표에서 더 좋아 보인다.

틀리는 모양 셋

1

없는 사실을 지어낸다

2

진짜 위에 가짜 한 줄

3

살짝 다른 이야기로 넘어간다

증상의 유형을 알아 두면 현장에서 알아채기 쉽다. 없는 사실의 창작(존재하지 않는 판례·논문을 인용), 그럴 듯한 세부의 오염(실존 인물에 가짜 이력 한 줄), 자신 없는 주제에서의 매끄러운 우회(질문과 살짝 다른 이야기를 유창하게)다.

문장이 자연스러워도 맞는다는 뜻은 아니다

공통점은 문장의 품질로는 구별할 수 없다는 것이다. 유창함은 사전학습이 보장하지만 사실성은 누구도 보장하지 않으므로, “글이 자연스러우니 맞겠지”라는 직관이 정확히 함정이 된다.

세 회사가 직접 하는 말

OpenAI

The model may express high confidence even in incorrect answers. — OpenAI Help, Does ChatGPT tell the truth?

틀린 답에도 자신감을 보일 수 있다

Google

Gemini can hallucinate and present inaccurate information as factual. — Google, Gemini Apps Help

지어낸 정보를 사실처럼 내놓을 수 있다

Anthropic

...while these techniques significantly reduce hallucinations, they don't eliminate them entirely. Always validate critical information — Anthropic, Reduce hallucinations

줄일 수는 있어도 없앨 수는 없다. 중요한 정보는 늘 확인하라

외부에서 확인할 것

인용

숫자

기술 정보

바깥 문서를 가리킨 말(출처)

Always verify quotes, data, technical information or references to external documents. — OpenAI Help,
Does ChatGPT tell the truth?

추천 도서 15권 가운데 10권이 없는 책

2025년 5월 · 미국 신문 - 시카고 선타임스

1 2025년 5월 18일, 미국 신문 시카고 선타임스의 여름 특집에 추천 도서 15권이 실렸다.

2 15권 가운데 10권은 세상에 없는 책이었다. 저자 이름은 실제 작가였다.

3 글을 쓴 외부 필자는 AI 를 쓰고 사실 확인을 하지 않았다고 인정했다.

Chicago Sun-Times 입장문(2025-05-20) · CEO 사과문(2025-05-29)

진짜 이름이 붙으면 가짜도 더 그럴듯하다



진짜 이름 옆의 가짜는 더 그럴듯하다



이름이 맞아도 그 옆의 제목과 숫자는 따로 확인한다

작가 한 명의 책 다섯 권을 물으면

소설가 김영하의 책 열 권을 제목과 처음 펴낸 해와 함께 알려 줘.

Gemini · Gemini 3.8 Flash

여기서 실제 화면으로 함께 본다

도서관 검색에 그 책이 있나

확인할 곳 국립중앙도서관 소장자료 검색 · 서점 누리집

1

빈칸 1

AI 의 답

2

빈칸 2

찾은 곳의 값

3

빈칸 3

같은가

이번 학기 중간고사 기간을 물으면

경기과학기술대학교 2026학년도 2학기 중간고사 기간이 언제부터 언제까지야?

Claude · Claude 4.6 Sonnet

여기서 실제 화면으로 함께 본다

학사 공지에서 직접 찾으면

확인할 곳

경기과학기술대학교 누리집 · 공지사항(교무팀)

1

AI 의 답

2

찾은 곳의 값

3

같은가

성인 카페인 하루 권장량을 물으면

우리나라 식품의약품안전처가 권하는 카페인 하루 최대 섭취량을 성인, 임산부, 어린이와 청소년으로 나눠 알려 줘.

ChatGPT · GPT-5.6 Sol

여기서 실제 화면으로 함께 본다

식약처 자료의 숫자와 대 보면

확인할 곳 · 식품의약품안전처 누리집

빈칸 1

AI 의 답

성인
임산부
어린이와 청소년



빈칸 2

찾은 곳의 값

식약처 자료에서
찾은 숫자와 단위



빈칸 3

같은가

숫자와 단위
「이하」라는
말까지

올해 최저임금을 물으면

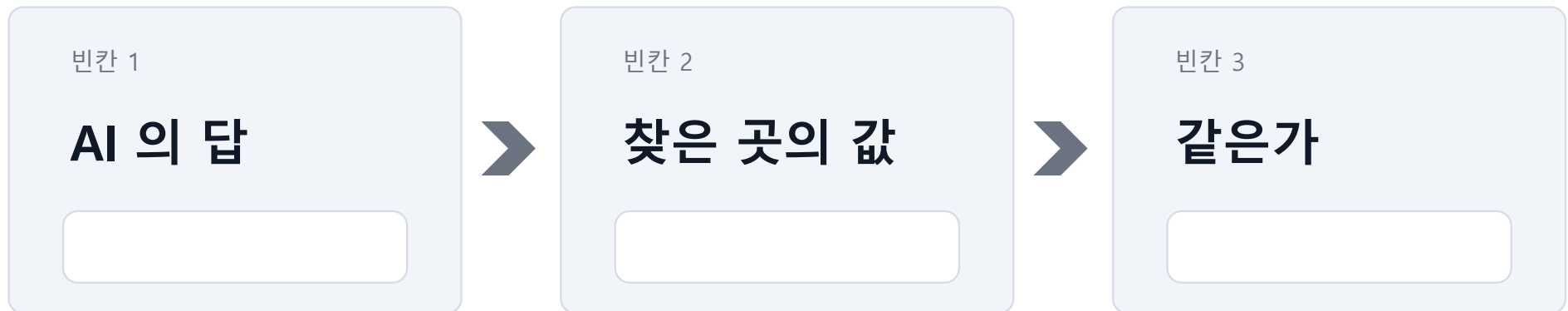
올해 우리나라 최저임금은 시간당 얼마야? 몇 년도 기준인지도 같이 말해 줘.

Gemini · Gemini 3.8 Flash

여기서 실제 화면으로 함께 본다

정부 고시와 대 보면, 시간이 멈춘 답

확인할 곳: 고용노동부 누리집 · 최저임금 고시



원인 셋째, 시간이 멈춰 있다.

AI 가 준 근거 주소를 직접 열어 보기

1

AI에 다음 장 프롬프트를
그대로 보낸다

2

받은 주소를 새 탭에서
하나씩 연다

3

열렸나, 같은 내용인가를
표에 적는다

AI 에게 근거 주소를 달라고 하면

만 6~18세 청소년에게 원하는 하루 신체활동 시간을 알려 줘.
그 근거가 되는 공식 기관 자료의 주소를 두 개 붙여 줘.

Claude · Claude 4.6 Sonnet

여기서 실제 화면으로 함께 본다

주소가 열리나, 열리면 같은 내용인가

1

주소가 열렸나

2

열렸다면 그 쪽에 AI
가 말한 숫자가 있나

3

기관 이름은 맞는데
쪽이 없는 주소가 있
었나

주소가 열리는 것과 내용이 맞는 것은 다른 일이다.

직접 실습

있는 책 없는 책 · 시연 작가

작가 한강의 책 다섯 권을 제목, 처음 펴낸 해, 출판사와 함께 알려 줘.

Claude · Claude 4.6 Sonnet

여기서 실제 화면으로 함께 본다

다섯 권 가운데 몇 권이 서점에 있었나

1

다섯 권 가운데 몇 권
이 서점에 있었나

2

제목은 있는데 펴낸 해
가 다른 책이 있었나

3

없는 책이 있었다면 저
자는 진짜였나

없는 과목을 물으면 · 두 번

아래 질문에 두 번 답해 줘.

질문: 경기과학기술대학교 교양 과목 「고양이 행동과 데이터」는 몇 학점이고 누가 가르쳐?

첫 번째: 그냥 답해 줘.

두 번째: 확실히 아는 것만 답하고, 모르면 모른다고 답해 줘.

Claude · Claude 4.6 Sonnet

여기서 실제 화면으로 함께 본다

「모르면 모른다고」한 줄이 바꾼 것

1

첫째 답이 없는 것을 있는 것처럼 소개했나

2

둘째 답은 모른다고 했나

3

한 줄을 붙여도 지어낸 화면이 있었나

이 동아리와 과목은 실제로 없다.

할인 합계, 계산기로 다시 세기

한 벌에 29,900원인 티셔츠 3벌을 20% 할인받아 사고, 배송비 3,000원을 더하면 모두 얼마야? 합계만 알려 줘.

Gemini · Gemini 3.8 Flash

여기서 실제 화면으로 함께 본다

직접 실습

나눠 내기 합계, 계산기로 다시 세기

치킨 2마리(한 마리 21,000원), 콜라 1병 2,500원, 배달 팁 3,000원을 다섯 명이 똑같이 나누면 한 사람이 얼마야? 한 사람 몫만 알려 줘.

ChatGPT · GPT-5.6 Sol

여기서 실제 화면으로 함께 본다

계산기 값과 같았나

1

계산기 값과 같았나

2

다르다면 어느 단계
에서 같았나

3

같았다면, 다음에도
다시 셀 것인가

좋은 모델로 바뀌어도 사라지지 않는다

할루시네이션은 버그가 아니라 다음 토큰 예측이라는 설계의 그림자이며, 따라서 “더 좋은 모델”만으로는 사라지지 않고 사용하는 쪽의 전략이 필요하다.



**자신 있게 틀리는 것은 고장이 아니다.
그래서 쓰는 쪽에 절차가 필요하다.**

2교시: 밖에서 확인하는 절차



2교시

직접 확인하는 절차

되묻지 말고, 밖에서 찾아 확인한다.

해결책 셋

네 동작

내 기획안 검증

AI 에게 「정말 맞아?」라고 되물으면

정말 맞아? 확실해?

ChatGPT · GPT-5.6 Sol

여기서 실제 화면으로 함께 본다

되묻기는 확인이 아니다

AI 에게 다시 묻기

- 같은 확률 게임을 한 번 더 한다
- 바뀌도 우겨도 근거는 없다

밖에서 찾기

- 학교 누리집, 기관 자료, 원문
- 찾은 곳을 적을 수 있다

확인은 AI 밖에서 한다.

확실한 해결책 셋

1

출처를 달라고
한다

2

원문을 먼저 준다

3

팩트 체커 역할
을 준다

셋 다 쓰고 나서도 확인은 내가 한다.

출처를 달라고 한다, 그리고 직접 연다

출처 요구하기: AI 의 답변에 의심이 들 때는 즉시 출처를 요구하자. "그 주장의 근거가 되는 실제 뉴스 기사나 논문의 URL 주소를 알려줘."라고 지시하면, AI 가 정보를 지어냈는지 혹은 실제 데이터에 기반한 것인지 확인할 수 있다. (물론, 이때 제시하는 URL 도 가짜일 수 있으므로 직접 확인은 필수.)

(물론, 이때 제시하는 URL 도 가짜일 수 있으므로 직접 확인은 필수.)

출처를 붙여 답하게 하면

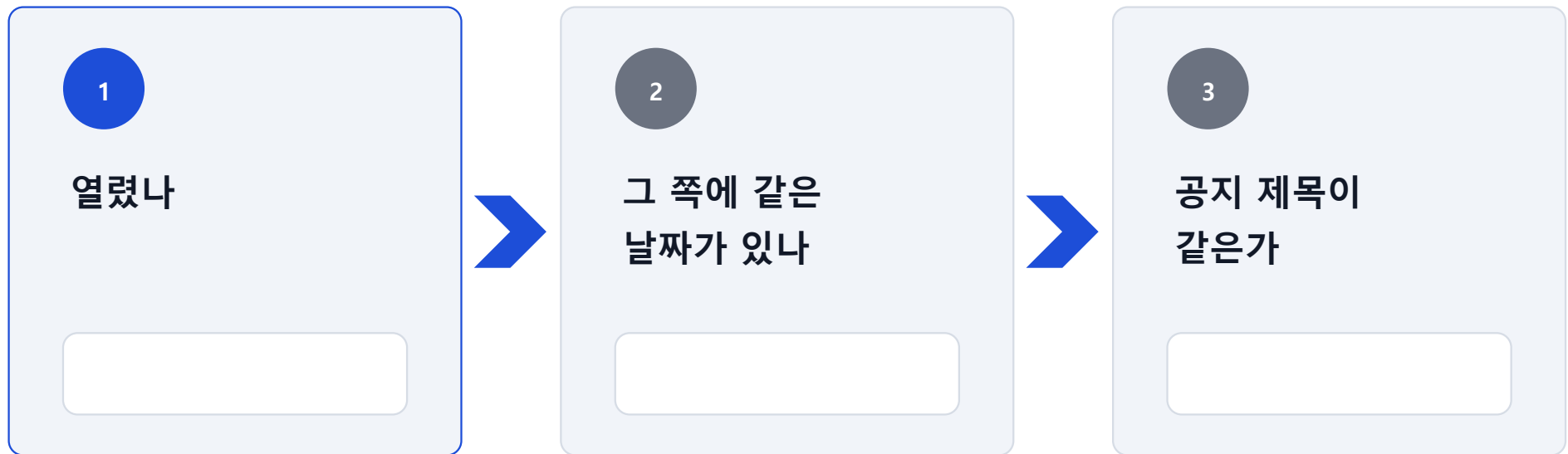
경기과학기술대학교 2026학년도 2학기 중간고사 기간을 알려 줘.
학교 공지의 제목과 주소도 같이 붙여 줘.

Claude · Claude 4.6 Sonnet

여기서 실제 화면으로 함께 본다

받은 주소를 하나씩 열어 본다

확인할 곳: AI 가 준 주소 · 학교 공지사항



원문을 먼저 주고 그 안에서만 답하게 한다

당신이 신뢰하는 기사나 보고서의 원문을 프롬프트에 먼저 제공(Context 법칙)하고, “오직 이 내용에만 근거해서 질문에 답해줘”라고 지시하도록 하자.

“아래 문서를 근거로 답하라. 근거가 없으면 없다고 답하라.”

4주차 맥락 블록 자리

「프롬프트」 p.239 · 「개론」 p.304

학교 공지 원문을 붙여 물으면

아래 공지에만 근거해서 답해 줘. 공지에 없으면 없다고 답해 줘. 1) 등록금납입증명서는 언제부터 출력할 수 있어? 2) 휴대폰 앱에서도 출력할 수 있어?

[공지]

등록금납입증명서는 등록금 납입일 다음날부터 출력이 가능합니다.

참고하시기 바랍니다.

1. 종합정보시스템 접속 (<https://portal.gtec.ac.kr>)
2. 로그인 (학번/비밀번호)
3. 인트라넷 클릭
4. 인트라넷-학생서비스-등록-등록금납입증명서출력 클릭

Gemini · Gemini 3.8 Flash

여기서 실제 화면으로 함께 본다

모른다고 말해도 된다고 적어 준다

Allow Claude to say "I don't know": Explicitly give Claude permission to admit uncertainty. This simple technique can drastically reduce false information.

모른다고 해도 된다고 분명히 허락하라. 이 간단한 방법이
지어낸 정보를 크게 줄인다.

빈칸도 0점이 아니라고 알려 주기

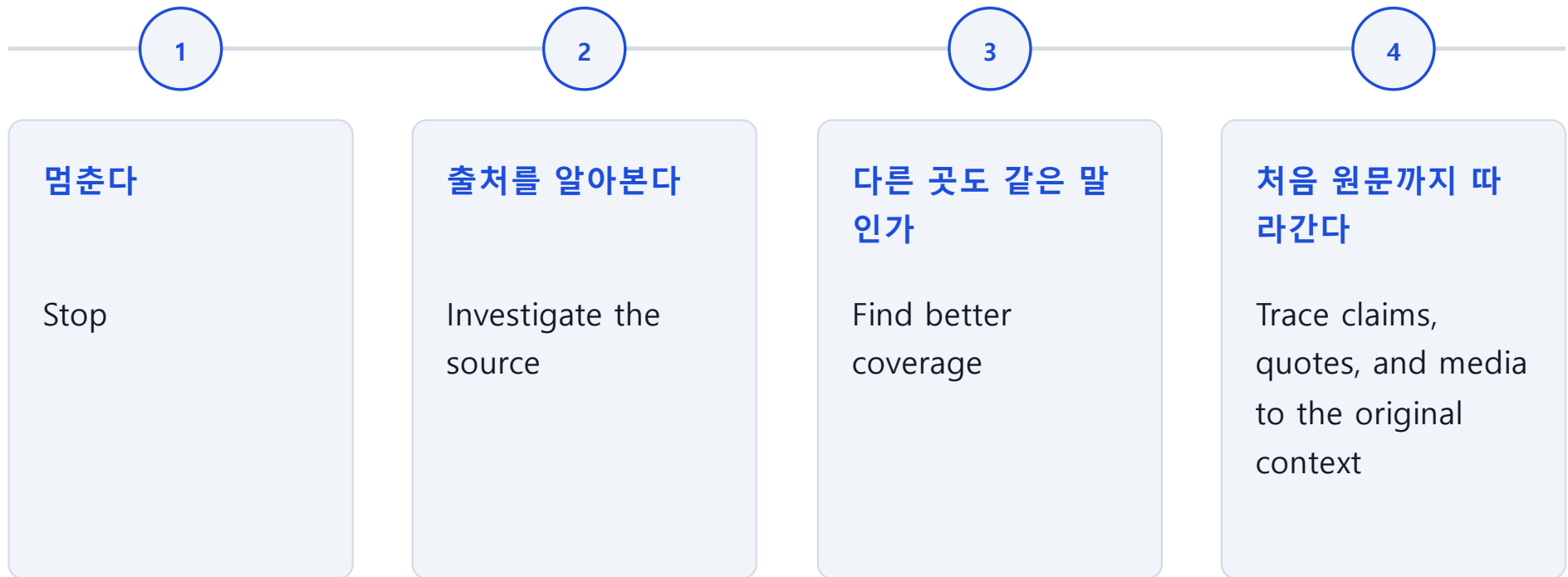
팩트 체커 역할을 줘도 확인은 내가 한다

‘팩트 체커’ 역할 부여하기: “너는 의심 많고 꼼꼼한 팩트 체커야. 너의 최우선 임무는 정확성이야. 100% 확실하지 않은 정보에 대해서는 ‘불확실함’이라고 명시하고, 절대로 추측해서 답변하지 마.” 와 같이 강력한 페르소나를 부여하여 답변의 신뢰도를 높일 수 있다.

「프롬프트」 p.239~240

역할은 답을 조심스럽게 만든다. 맞는지는 여전히 밖에서 본다.

밖에서 확인하는 네 동작



Mike Caulfield, SIFT (2019)

멈춘다

그럴듯한 답

그럴듯하면 바로 쓰고 싶어진다



쓰기 전에

쓰기 전에 「이건 어디서 나온 말이지?
」를 먼저 묻는다

출처를 알아본다

1

누가 말했나

그 말을 한 곳이 누구인가(기관, 신문, 개인)

2

자격이 있나

그 곳이 그 일을 말할 자격이 있는가(학교 일은 학교 누리집)

카페인은 식약처 · 최저임금은 고용노동부 · 학사 일정은 학교

다른 곳에서도 같은 말을 하나 찾는다

한 곳

한 곳만 보고 끝내지 않는다

서로 다른 두 곳

서로 다른 두 곳이 같은 값을 말하면 믿을 만하다

처음 나온 원문까지 따라간다



요약을 거치는 동안 숫자와 낱말이 바뀌기도 한다.

전문가는 새 탭을 열었다

2019년 미국 스탠퍼드대 연구진이 사실검증 전문가 10명, 역사학 박사 10명, 대학생 25명에게 낯선 누리집의 믿을 만함을 판단하게 했다.

사실검증 전문가들은 그 누리집 안에서 오래 읽지 않고 새 탭을 열어 바깥에서 그 출처를 확인했고, 더 짧은 시간에 더 타당한 결론을 냈다.



안에서 오래 읽기보다 밖에서 짧게 확인한다.

Wineburg-McGrew, Teachers College Record (2019)

5년 동안 살아남은 가짜 전쟁

2007~2012 · 영어 위키백과

After careful consideration and some research, I have come to the conclusion that this article is a hoax—a clever and elaborate hoax, but a hoax nonetheless. An online search for "Bicholim conflict" or for many of the article's purported sources produces only results that can be traced back to the article itself. Take, for example, one of the article's major sources: Thompson, Mark, *Mistrust between states*, Oxford University Press, London 1996. No record at WorldCat. No mention at the OUP site. No used listings at Alibris or ABE. I can find no evidence anywhere that this book exists.

2007년 7월, 영어 위키백과에 17세기 인도에서 있었다는 전쟁 「Bicholim conflict」 문서가 올라왔다. 같은 해 우수 문서로도 뽑혔다.

2012년 12월, 편집자 한 사람이 문서가 인용한 책을 하나씩 찾아보다가 그 책들이 도서관 목록 어디에도 없다는 것을 밝혀냈다.

문서는 5년여 만에 지워졌다. 삭제 토론의 결론은 「A snow job for a very devious hoax.」였다.

인용된 책을 하나씩 찾아본 사람이 잡았다

99

출처가 붙어 있다고 출처가
있는 것은 아니다



인용된 책과 글이 실제로 있
는지 하나씩 찾는다

숫자는 다시 센다



합계, 나눗셈, 할인은 계산
기로 다시 센다



단위(원, 퍼센트, mg)까지
같이 본다

풀이가 보이면 어느 단계에서 틀렸는지 찾는다

비슷한 것이 나왔다고 같은 것은 아니다



검색에 비슷한 제목이 나왔
다



제목, 저자, 연도가 모두 같
다

셋을 하나씩 대 보고 나서야 「있다」고 적는다.

맞음, 틀림, 확인 안 됨

맞음

그대로 쓴다

틀림

찾은 값으로 고친다

확인 안 됨

쓰지 않거나 확인 못
함이라고 적는다

확인 안 됨을 맞음으로 세지 않는다.

「결과를 확인할 수 있는가」

LOOKING BACK

Can you check the result? Can you check the argument?

되돌아보기 · 결과를 확인할 수 있는가? 풀이를 확인할 수 있는가?

Pólya, How to Solve It (1945)

교수 시연

직접 실습

명언 주인 찾기 · 시연 명언

「천재는 1%의 영감과 99%의 노력으로 이루어진다」는 말은 누가 했고, 처음 기록된 것은 언제 어느 때
체야?

1

AI 가 댄 사람과 처음 출처가 같았나

2

처음 기록의 해와 매체를 찾았나

3

유명한 사람 이름이 붙으면 왜 더 믿게 될까

링크 열어 보기 · 시연 주제

한글날이 다시 공휴일이 된 해를 알려 줘.
그 근거가 되는 자료의 주소를 두 개 붙여 줘.

Gemini · Gemini 3.8 Flash

여기서 실제 화면으로 함께 본다

직접 실습

링크 열어 보기 · 다른 주제

어린이날이 공휴일이 된 해를 알려 줘.
그 근거가 되는 자료의 주소를 두 개 붙여 줘.

ChatGPT · GPT-5.6 Sol

여기서 실제 화면으로 함께 본다

학교 도서관이 토요일 몇 시까지 여는지 물으면

경기과학기술대학교 도서관은 학기 중 토요일에 몇시부터 몇시까지 열어?

Claude · Claude 4.6 Sonnet

여기서 실제 화면으로 함께 본다

도서관 누리집 이용 시간표와 대 보면

확인할 곳 · 경기과학기술대학교 도서관 이용 안내

빈칸 1

토요일 운영

()

빈칸 2

닫는 시간

()

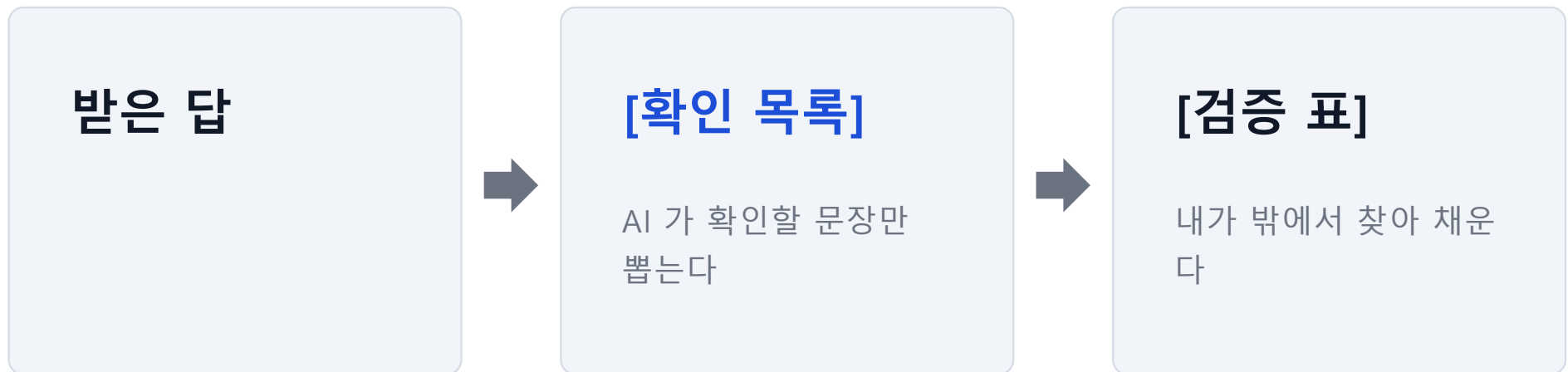
빈칸 3

휴관일

()

질문 자체가 틀렸을 수도 있다.

뽑기는 AI 가, 확인은 내가



AI 에게는 판단을 맡기지 않는다.

[확인 목록]

[확인 목록]

[받은 답]

(AI 가 준 답을 그대로 붙인다)

위 답에서 사실을 말하는 문장만 번호를 붙여 뽑아 줘. 고치거나 판단하지는 마.

문장마다 숫자, 날짜, 이름, 출처 가운데 무엇이 들어 있는지 표시해 줘.

[검증 표]

| 번호 | 확인할 주장 | 확인한 곳(주소나 자료 이름) | 결과(맞음 · 틀림 · 확인 안 됨) | 고친 내용 |

번호	확인할 주장	확인한 곳(주소나 자료 이름)	결과(맞음 · 틀림 · 확인 안 됨)	고친 내용

6주차 기획안 뒤에 붙는 자리

6주차 네 단계와 받은 기획안

[확인 목록] (오늘)

[검증 표] (오늘, 내가 채움)

지난 양식은 바꾸지 않고, 받은 결과 뒤에 붙인다.

2교시

이제부터는 직접 실습



6주차 기획안과 검증 표 빈 양식 준비하기

내 6주차 기획안 꺼내기

1

6주차 기획안을 열고 [확인 목록]으로 확인할 문장을 뽑는다

2

[검증 표]를 직접 검색해 채운다

3

틀린 것을 고쳐 기획안을 다시 받는다

제출: 확인 목록, 채운 검증 표, 고친 기획안 (LMS)

기획안에서 확인 목록 뽑기

[확인 목록]

[받은 답]

(AI 가 준 답을 그대로 붙인다)

위 답에서 사실을 말하는 문장만 번호를 붙여 뽑아 줘. 고치거나 판단하지는 마.
문장마다 숫자, 날짜, 이름, 출처 가운데 무엇이 들어 있는지 표시해 줘.

Claude · Claude 4.6 Sonnet

여기서 실제 화면으로 함께 본다

검증 표 채우기 · 직접 검색

확인 목록의 번호마다 한 줄씩, 밖에서 찾아 채운다.

[검증 표]

| 번호 | 확인할 주장 | 확인한 곳(주소나 자료 이름) | 결과(맞음 · 틀림 · 확인 안 됨) | 고친 내용 |

번호	확인할 주장	확인한 곳	결과	고친 내용

교수 예시 · 채운 검증 표 한 장

| 번호 | 확인할 주장 | 확인한 곳(주소나 자료 이름) | 결과(맞음 · 틀림 · 확인 안 됨) | 고친 내용 |

| 1 | 중간고사는 2026년 10월 19일~24일 | 학교 공지사항 「2026-2학기 중간고사 안내」 | 맞음 | — |

| 2 | 토요일 오후 도서관 C-Zone 에서 모인다 | 도서관 누리집 이용 안내 | 틀림 | 토·일요일은 휴관. 금요일 오후로 바꾼다 |

| 3 | 스터디룸은 앱으로 예약한다 | 도서관 누리집에서 찾지 못함 | 확인 안 됨 | 확인 못 함. 도서관에 직접 묻는다 |

틀린 것을 고쳐 기획안 다시 받기

[앞 단계 결과]
(내 기획안을 붙인다)

[검증 표]
(채운 검증 표를 붙인다)

위 검증 표에서 결과가 「틀림」인 줄만 「고친 내용」 값으로 바꿔 기획안을 다시 써 줘.
「확인 안 됨」인 줄은 지우거나 「확인 못 함」이라고 적어 줘. 다른 곳은 바꾸지 마.

여기서 실제 화면으로 함께 본다

무엇이 틀렸고 무엇을 확인 못 했나

1

확인한 곳 칸이 빈 줄
이 있나

2

확인 안 됨을 맞음으
로 센 줄이 있나

3

틀림이 하나라도 나
왔나

점검은 안에서, 검증은 밖에서

점검은 계획 안에서 서로 맞는지,

검증은 밖에서 실제와 맞는지.

안 될 때 여기부터

틀림

- AI 에게 「맞아?」라고 되물음
- 주소만 보고 열어 보지 않음
- 확인 안 됨을 맞음으로 셈

맞음

- ✓ 밖에서 직접 찾기
- ✓ 주소를 열어 내용까지 읽기
- ✓ 확인 안 됨은 따로 적기

뽑기는 AI 가, 확인은 내가

오늘 공부한 내용

환각은 그럴듯한 말로 자신 있게 틀리는 것

글이 자연스러워도 맞다는 뜻은 아니다. 밖에서 확인한다

원인 셋: 확률로 고른다 · 모르는 줄 모른다 · 시간이 멈췄다

멈추고, 출처를 보고, 다른 곳도 보고, 원문까지

[확인 목록]은 AI 가, [검증 표]는 내가

다음 주 중간 고사 진행

문제 정의부터 검증까지 지금까지 배운 것을 혼자 해 봅니다



다음 주 예고

- 중간 고사