

SOCIAI · 무료 교육자료 · 90분 워크숍

내 업무로 만드는 AI 평가 실습

합성 회의록을 바탕으로, 실제 업무에 맞는 조건을 세우고 기록하는 방법을 익히는 인쇄용 교안입니다.

이 교안의 회의록과 답변은 수업을 위해 만든 합성 예시이며 실제 모델의 측정 결과가 아닙니다. 결과를 실제 업무에 적용하기 전에는 다른 입력과 실패하기 쉬운 사례에서도 확인하세요.

이 수업을 마치면

- 반복 업무 하나를 고르고, 반드시 맞아야 하는 조건과 멈춰야 할 실패를 설명할 수 있습니다.
- 응답 완료, 출력 형식, 내용 정확성, 기다린 시간을 나눠 기록할 수 있습니다.
- 자동 채점의 오답을 찾고, 다음 비교 실험에 남길 입력과 조건을 정할 수 있습니다.

남길 결과물: 업무 평가표 1장, 같은 입력의 답변 비교 기록, 채점 반례 1개, 다음에 바꿀 조건 1개.

수업에 넣는 방법

기존 과정	넣을 위치	활동과 결과물
AI 문해력·입문	첫 요약·문서 작성 실습 뒤 20분	아래 합성 답변 두 개를 채점하고, 그럴듯함과 정확성의 차이 찾기
업무 활용·프롬프트	업무 템플릿을 만든 뒤 45분	같은 입력에 현재·수정 프롬프트를 적용하고 평가표 작성
에이전트·조직 도입	자동화 파일럿을 고를 때 90분	전체 실습 후 실행 실패·복구·사람 검토 기준까지 정하기

20분 활동은 교안 속 답변으로 진행하므로 새 계정이나 유료 모델이 없어도 됩니다. 실제 AI를 사용하는 활동에서는 이미 이용 가능한 도구를 쓰고 민감한 업무 원문 대신 아래 합성 입력으로 시작합니다.

90분 진행 순서

시간	참가자 활동	진행자가 확인할 것
0-10분	말기고 싶은 업무 한 가지와 잘못됐을 때의 영향을 적기	여러 업무를 한꺼번에 평가하지 않는가
10-25분	합성 회의록을 읽고 필수 조건 세 개 정하기	정답과 판단 보류를 구분했는가
25-40분	합성 답변 A·B를 각각 채점하고 짝과 비교	단어의 존재와 담당자·날짜의 연결을 함께 봤는가
40-60분	같은 입력으로 현재·후보 프롬프트를 각 3회 실행	새 대화로 시작하고 한 번에 한 조건만 바꿨는가
60-75분	완료·형식·내용·시간을 따로 정리	실패·미실행·판단 보류를 빠뜨리지 않았는가
75-90분	답변 B의 셀 하나를 바꾼 오답을 만들고 동료의 규칙으로 검사, 다음 실험 정하기	반례를 잡아냈는가, 제한된 결과를 과도하게 일반화하지 않는가

모델 응답이 늦으면 시간 상한에 도달한 실행을 기록하고 합성 답변 토론으로 이어갑니다. 미실행 칸은 비워두거나 미실행으로 표시합니다. 시간을 맞추려고 성공한 결과만 골라 넣지 않습니다.

참가자 실습: 회의록의 마지막 결정 찾기

다음 블록 전체를 복사합니다. 날짜와 인물은 교육용 가상 정보입니다.

[교육용 가상 회의록 v1]

2026년 9월 15일 회의.

- 안내문 초안은 민지가 9월 18일까지 작성하기로 했다.
- 신청 폼 점검은 준호가 맡는다. 마감일은 정하지 않았다.
- 회의 마지막에 안내문 초안 담당자를 서연으로 변경했다.
안내문 초안 마감일은 9월 18일로 유지했다.

[요청]

마지막 결정 기준으로 업무별 담당자와 마감일을 정리해 주세요.

열 이름은 '업무 | 담당자 | 마감일'인 표를 사용하세요.

정해지지 않은 값은 '미정'이라고 적으세요.

실행 전에 합의할 필수 조건

- 안내문 초안의 담당자는 서연이며 마감일은 2026-09-18입니다.
- 신청 폼 점검의 담당자는 준호이며 마감일은 미정입니다.
- 두 업무가 빠짐없이 들어 있고 각각의 담당자·날짜가 올바른 업무에 연결됩니다.

별도의 형식 조건: '업무 | 담당자 | 마감일' 세 열을 가진 표입니다. 내용은 맞지만 표 형식을 어긴 답변이라면 내용과 형식을 각각 판정합니다.

의미가 같은 날짜 표기(예: 9월 18일)는 내용 판정에서 허용합니다. 정확한 YYYY-MM-DD 표기가 업무에 필요하다면 실행 전에 프롬프트와 형식 조건에 명시합니다. 결과를 본 뒤 특정 답변을 통과시키기 위해 기준을 바꾸지 않습니다.

AI를 켜기 전에: 채점자의 시험 문제

아래 답변은 비교 토론용으로 사람이 만든 예시입니다.

답변 A

업무	담당자	마감일
안내문 초안	민지	2026-09-18
신청 폼 점검	준호	2026-09-19

답변 B

업무	담당자	마감일
안내문 초안	서연	2026-09-18
신청 폼 점검	준호	미정

진행자 토론 질문

- 어떤 셀이 틀렸고, 회의록 어느 문장이 그 판단의 근거인가요?
- 형식 통과와 내용 통과를 한 점수로 합치면 무엇이 가려지나요?
- 마감일을 '미정'으로 남긴 답변을 불완전하다고 감점해도 될까요?

현재 설정과 후보 설정 비교하기

첫 비교는 같은 모델에서 프롬프트 한 가지를 바꾸는 방식으로 시작할 수 있습니다. 현재 요청을 기준으로 두고, 후보에는 '변경·취소 사항을 먼저 반영한 뒤 최종 표를 작성해 주세요'를 추가합니다. 이것은 시험할 후보이며 개선이 보장된 문구는 아닙니다.

각 설정을 새 대화에서 세 번씩 실행합니다. 현재→후보, 후보→현재처럼 순서를 번갈아 쓰고 같은 입력을 유지합니다. 요청 시각부터 최종 답변까지 시간을 재며, UI에 표시된 모델명과 확인할 수 있는 설정을 적습니다. 확인하지 못한 설정은 '미확인'으로 남깁니다.

같은 입력을 세 번 실행한 것은 독립적인 업무 세 개가 아닙니다. 세 번 모두 통과해도 이후 업무의 신뢰도 100%를 입증하지 않습니다. 다른 입력과 실패하기 쉬운 사례를 추가하는 출발점으로 사용합니다.

기록표와 판정 기준

구분	확인 질문	기록값
실행	정한 시간 안에 최종 답변이 왔나요?	완료 / 시간 초과 / 오류 / 미실행
형식	요청한 표·열이 있나요?	통과 / 실패 / 미확인
내용	모든 필수 조건과 원문 근거가 맞나요?	통과 / 실패 / 판단 보류
시간	요청부터 답변 완료까지 얼마나 걸렸나요?	초, 미완료면 관찰한 상한 시간
사람의 수정	그대로 쓸 수 있나요?	수정할 셀·이유·확인할 정보

미완료 응답은 내용 통과로 세지 않습니다. 전체 시도 수와 완료 수를 먼저 적고, 내용 판정을 마친 답변 중 통과 수를 따로 적습니다. 예를 들어 6회 시도 중 완료 5회, 그중 내용 통과 4회라면 두 분모를 모두 보여줍니다. 이 숫자도 설명용 예시입니다.

진행자의 결과 검토표

아래는 참가자가 만든 평가 설계의 검토 기준입니다. AI 모델 성능 점수와 구분합니다.

- **입력과 기준:** 동일한 입력이 보존되고, 담당자 변경과 미정 날짜를 다루는 조건이 실행 전에 적혀 있습니다.
- **기록과 근거:** 실패·미실행·판단 보류가 남아 있고, 내용 판정에 원문의 구체적인 근거가 붙어 있습니다.
- **해석과 다음 행동:** 반복 횟수의 한계를 설명하며, 아직 시험하지 않은 업무와 다음에 바꿀 변수 한 가지를 제시합니다.

세 항목이 갖춰지면 '평가 설계 초안 완성'으로 봅니다. 모델을 업무에 자동 투입해도 된다는 승인은 아닙니다. 동료와 판정이 다르면 불일치 이유부터 기록하고 해당 기준을 보완합니다.

다음 수업: 내 프로젝트에 옮기기

회의록 예제를 실제 업무의 합성·비식별 사례로 바꿉니다. 담당자 변경, 취소, 누락처럼 틀리기 쉬운 입력과 정상 입력을 함께 넣습니다. 수정에 쓰지 않은 별도 사례도 남겨둡니다.

개발자는 입력 스냅샷과 프롬프트 버전, 요청 모델과 실제 응답 모델, 대체 경로, 시간 초과·형식 오류·내용 오류를 기록합니다. 사람은 '내용이 맞는가'와 '독자에게 이 행동을 권해도 되는가'를 각각 검토합니다.

- [입문 가이드와 작성형 평가표](#)
- [하네스·파이프라인 적용 가이드](#)
- [배경 기사](#)
- [실습 기록 남기기](#)

foundby 공개 자료에서 업무별 필수 조건, 판정 방식, 반복 실행, 실패 분류의 설계를 참고했습니다. 원문의 모델 성능 결과를 재현한 수업은 아닙니다. 회의록·답변 A/B·시간표·검토표는 Sociai가 새로 구성했습니다.

- [foundby 평가 방법](#)
- [foundby 공개 테스트 케이스](#)
- [foundby 운영 사례와 최적화 검증](#)

원문 검토일: 2026-09-14. 교안 공개일: 2026-09-15.

참가자용 빈 결과 기록표

수업 중 정한 조건과 실행 결과를 적으세요. 빈 칸은 억지로 채우지 않아도 됩니다.

내 업무 / 입력 버전	필수 조건 / 치명 실패

현재 설정 / 비교할 설정 / 기다릴 시간

반복	모델·설정	실행	형식	내용	시간	근거 간단
1						
2						
3						

참가자용 빈 결과 기록표 (계속)

반복	모델·설정	실행	형식	내용	시간	근거 간단
4						
5						
6						

다음에 바꿀 조건 / 다른 입력에서 확인할 점