

AI/ML 연재 공장관리자의 첫 AI 실습 9/12

AI 답의 오류, 내 불량 기록으로 10분 안에 잡기 (40분)

방동걸 (주)웨이스 대표이사

2026년 9월 21일

오늘은 우리 공장용 「검증 체크표 5문」 1장을 만들고, 앞서 만든 결과물 2개를 그 체크표로 다시 검사합니다. 40분이면 됩니다. 제조 AI 운영체제 VEXPLOR 를 만드는 웨이스가 현장에서 자주 듣는 말이 「AI 가 근거까지 댔는데 열어 보니 아니더라」입니다. 근거를 댔 답을 10분 안에 대조하는 순서가 오늘 만드는 것입니다.

오늘 새로 배우는 것 하나: 근거의 모양과 근거는 다르다. 「3행에 따르면」은 대조할 자리를 알려 줄 뿐, 맞다는 뜻이 아닙니다. 왜 범용 AI 를 제조 현장에서 그대로 믿기 어려운지는 앞선 설명 글 제조 현장에서 ChatGPT를 못 믿는 이유, 그리고 환각 없는 제조 AI에 있으므로 여기서는 이유를 반복하지 않고 손으로 잡는 절차만 다룹니다.

연재 공통 규칙은 이번에도 같습니다.

- 붙여넣기 전에 가립니다 — 고객사명·단가·사람 이름·도면.
- AI 가 낸 숫자는 엑셀에서 같은 숫자를 다시 냅니다. 다르면 AI 가 틀린 것입니다.
- 지시문 끝에 「빈 셀은 빈 셀로, 모르면 [확인 필요]로 적어라」를 붙입니다.
- 1편에서 만든 공장 설명 카드를 대화 첫 줄에 붙입니다.
- 결과물은 파일로 남깁니다 — 09_검증체크표_20260921.xlsx.

01

준비물은 앞 편 결과물과 불량 기록 엑셀입니다

항목	어디서	비용·조건
4편 불량 기록 엑셀	4편에서 저장한 파일	400행 안팎
6편 결과 표와 AI 답	6편 대화 기록	원인 문장 포함
AI 대화 도구	파일 첨부가 되는 것	무료판이면 학습 끄기
엑셀	피벗 기능	피벗 = 엑셀이 대신 세어 주는 표

자기 데이터로 하시면 됩니다. 예제가 필요하면 아래 표 10행을 엑셀에 붙여 넣어 시작하십시오. 예제 공장의 8월 불량 기록 400행 중 앞 10행이며, A 열에 1부터 번호를 넣어 두었습니다. 이 번호가 오늘의 핵심 도구입니다 — AI 가 「몇 행」이라고 말하면 이 번호로 찾습니다.

번호	불량명	수량	조
1	굽힘	12	주간
2	도금	3	야간
3	치수 불량	1	주간
4	찍힘	2	주간
5	스크래치	1	주간
6	굽힘	8	주간
7	버	1	야간
8	치수 불량	5	주간
9	굽힘	7	야간
10	찍힘	1	주간

붙여넣기로 할 때는 200행 안팎으로 잘라 넣습니다. 길게 넣을수록 뒤에서 확인할 함정이 커집니다.

02

순서는 시험 4개를 먼저 하고 체크표를 마지막에 확정합니다

체크표를 먼저 쓰고 시험하는 것이 아니라, 시험 4개를 직접 해 본 뒤 그 경험으로 5문을 확정합니다. 남이 준 체크표는 남지 않기 때문입니다. 아래 AI 답 예시는 예제 데이터에 혼한 오류를 심어 만든 것이라, 자기 도구에서는 다른 자리에서 틀립니다 — 그래서 절차가 필요합니다.

1단계 — 한 셀을 일부러 10배로 고쳐 넣습니다. 4편 엑셀 사본을 만들고 1행 굽힘 수량 12 를 120 으로 고칩니다. 그 사본을 붙여 넣고 4편과 같은 질문을 합니다.

공장 설명 카드 (붙여넣기) 아래 불량 기록에서 불량명별 수량 합계와 전체 합계를 표로 만들어라. 표기를 합치거나 나누지 말고 있는 그대로 세어라. 빈 셀은 빈 셀로, 모르면 [확인 필요]로 적어라.

예제 데이터로 이 시험을 하면 AI 초안은 이런 모양으로 옵니다.

전체 2,597개. 찍힘 549 · 굵힘 493 · 버 230 · 놀림 207 · 치수불량 187. 굵힘이 2위입니다.

숫자는 고친 데이터 기준으로 맞습니다. 그런데 이 파일에서 한 행의 수량은 최대 12개인데 120개짜리 행이 들어와도 아무 말이 없습니다. 여기서 「AI 는 시키지 않은 것은 보지 않는다」가 보이면 맞게 온 것입니다. 이상값(다른 행과 동떨어진 값)을 잡게 하려면 지시문에 한 줄을 더 붙입니다 — 「한 행의 수량이 다른 행의 5배를 넘으면 그 행 번호를 따로 적어라」. 그래야 1행이 표시됩니다.

2단계 — 같은 질문을 새 대화에서 다시 합니다. 사본이 아닌 원본 400행으로 1단계와 같은 지시문을 2번, 매번 새 대화에서 넣습니다. 예제 데이터의 초안 예시는 첫 답이 전체 2,489개, 두 번째 답이 2,412개입니다. 엑셀 SUM(합계 함수)은 2,489 입니다. 두 번째 답은 뒤쪽 행 일부를 빠뜨리고 센 것입니다. 「같은 질문 두 번에 답이 다르면, 어느 쪽이 맞느냐가 아니라 AI 가 세지 않고 어렵하고 있다는 신호다」가 보이면 맞게 온 것입니다.

시험	AI 답	엑셀 값	판정
10배 셀 넣기	2,597개, 경고 없음	2,597개	계산 맞춤·이상값 놓침
같은 질문 1회	2,489개	2,489개	맞음
같은 질문 2회	2,412개	2,489개	77개 빠짐

3단계 — 「근거를 대라」고 한 뒤 그 근거를 엮습니다. 원본 400행을 넣고 이렇게 묻습니다.

주간조와 야간조의 불량 수량을 비교하고, 근거가 되는 행 번호를 대라.

예제 데이터의 초안 예시는 이렇습니다.

3행(야간, 찍힘 9개)과 17행(야간, 버 25개)을 보면 야간조 불량률 주간조의 2.3배입니다.

행 번호가 있으니 믿고 싶어집니다. 번호 열로 찾아가면 3행은 주간조, 치수 불량 1개입니다. 17행은 야간조가 맞지만 burr 11개이지 25개가 아닙니다. 피벗으로 조별 합계를 내면 주간 1,484개, 야간 1,005개 — 야간이 주간의 0.68배입니다. 2.3배가 아니라 야간이 더 적습니다. 근거의 모양은 갖췄지만 근거가 아니었습니다.

근거	AI 가 말한 것	원본	판정
3행	야간, 찍힘 9개	주간, 치수 불량 1개	조·불량명·수량 다름
17행	야간, 버 25개	야간, burr 11개	수량 다름
조별 비교	야간이 2.3배	야간이 0.68배	방향이 반대

4단계 — 6편에서 AI 가 낸 「원인」을 데이터와 대조합니다. 6편 대화에서 AI 가 「주요 원인은 자재 입고 지연과 금형 교체 병목」이라고 했다면, 그 문장의 낱말이 수주·출하 파일의 열 이름에 있는지 봅니다. 예제 파일의 열은 7개입니다 — 수주번호·수주일·품목·수량·약속납기·출하일·고객구분. 「자재」도 「금형」도 열에 없습니다. 원인 문장 옆에 [데이터 밖]이라고 적습니다. 데이터에 없는 원인은 틀린 것이 아니라 확인되지 않은 것이고, 확인은 담당자에게 묻는 일입니다.

5단계 — 시험 4개의 경험으로 체크표 5문을 확정합니다. 아래 표를 자기 말로 고쳐 A4 한 장에 인쇄해 모니터 옆에 둡니다.

번호	질문	어떻게 보나	오늘 잡은 것
1	합계가 엑셀과 같은가	SUM·피벗	2단계 77개
2	근거 행이 원문과 같은가	번호 열로 열어 본다	3단계 3행·17행
3	모르는 것을 모른다고 했나	[확인 필요] 개수	1단계 이상값
4	가정을 밝혔나	표 밖 가정 문장	4단계 원인
5	같은 질문 두 번에 같은 답인가	새 대화로 재질문	2단계

6단계 — 이후 결과물마다 통과 표시를 남깁니다. 파일 이름 끝이나 첫 시트에 「체크표 5/5 통과, 날짜」를 적습니다. 통과 못 한 문항 번호도 적습니다 — 「4번 미통과, 원인 문장은 담당자 확인 대기」처럼.



근거 있는 답을 10분 안에 대조하는 순서 6단계

03

막히는 곳은 이 3가지입니다

행 번호가 서로 다릅니다. AI 가 말한 3행이 엑셀의 3행과 다릅니다. 원인은 엑셀은 머리글이 1행이고 AI 는 머리글을 뺀 순서로 세거나 그 반대이기 때문입니다. 해결은 A 열에 번호를 직접 넣고 「행 번호는 A 열의 번호로 말하라」고 지시하는 것입니다. 오늘 준비물 표에 번호 열을 둔 이유입니다.

10배 셀을 넣었는데 AI 가 알아챘습니다. 이번에는 잡았다고 안심하면 다음에 놓칩니다. 원인은 AI 가 매번 같은 기준으로 보지 않기 때문입니다. 해결은 잡든 못 잡든 「5배 넘는 행은 따로 적어라」를 지시문에 고정하는 것입니다. 우연히 잡은 것과 시켜서 잡은 것은 다릅니다.

같은 질문 두 번에 답이 같아서 넘어갔습니다. 예제 데이터에서도 2번 다 2,489개가 나올 수 있습니다. 같다는 것은 맞았다는 뜻이 아니라 「이번엔 같았다」는 뜻입니다. 해결은 5번 문항을 통과했어도 1번 문항(엑셀 SUM)을 건너뛰지 않는 것입니다. 5문은 서로를 대신하지 않습니다.

04

결과가 맞는지는 체크표가 무엇을 잡았는지로 봅니다

체크표 5문이 각각 오늘 시험에서 실제로 무엇을 잡았는지 마지막 열에 적혀 있으면 됩니다. 한 문항도 잡은 것이 없다면 그 문항은 아직 우리 공장의 문항이 아니므로 다음 편 결과물에서 다시 시험합니다.

여기 나온 값(2,489개·0.68배·77개 차이)은 모두 예제 데이터 기준입니다. 자기 데이터에서는 값이 다르고, 틀리는 자리도 다릅니다. AI 가 그럴듯하게 틀리는 지점은 오늘 3단계에 있습니다 — 행 번호·출처·비율까지 갖춘 문장이 원본과 다릅니다. 근거를 댔다는 것은 맞다는 뜻이 아니라 대조할 것이 생겼다는 뜻입니다.

근거의 모양과 근거는 다르다

시험 4개로 체크표 5문을 확정하고

근거는 원본에서 연다



이 글을 한 장으로 요약하면 이렇습니다 — 시험 4개로 체크표 5문을 확정하고, 근거는 원본에서 연다

05

오늘 만든 것은 체크표 1장과 재검증 기록 2건입니다

오늘 만든 것은 우리 공장용 검증 체크표 5문 1장, 그리고 4편 불량 기록과 6편 원인 문장을 다시 검사한 기록 2건입니다. 앞으로 AI 가 낸 표마다 이 5문을 적용하면, 틀린 답을 10분 안에 잡을 수 있게 됩니다. 결과 표나 사진을 wace2301@gmail.com 으로 보내 주시면 다음 편에 반영합니다.

WACE(웨이스)는 공장이 스스로 판단하고 멈추지 않게 하는 제조 AI 운영체제 VEXPLOR를 만드는 회사다.



원문 blog.vexplor.com/post/catch-ai-mistake-in-ten-minutes
발행 2026년 9월 21일 · 방동걸 (주)웨이스 대표이사

회사 주식회사 웨이스 (WACE Inc.) · 대표이사 방동걸 ·
사업자등록번호 439-87-02852
주소 (18462) 경기 화성시 동탄구 동탄영천로 150, 에이동(A동)
1702:1706호 (영천동, 실리콘밸리동탄)
연락처 전화 070-4730-0629 · contact@wace.me
홈페이지 wace.me · linkedin.com/company/wace-inc



원문



홈페이지



링크드인