AI/ML

상관은 알람을 만들고, 인과는 결정을 만든다

방동걸 (주)웨이스 대표이사

2026년 8월 19일

공장에 AI를 붙이면 대개 알람이 먼저 늘어납니다. 온도가 임계를 넘었다, 진동 패턴이 평소와 다르다, 불량률이 상승 추세다. 그런데 알람이 늘어난 만큼 결정이 빨라졌느냐고 물으면 대답이 잘 나오지 않습니다. 제조 현장에 VEXPLOR를 넣는 웨이스가 그 질문을 스스로에게 던지며 정리했습니다.

원인은 모델 성능이 아니라 질문의 종류에 있습니다. 알람은 "무엇과 무엇이 같이 움직이는가"에 답하면 만들 수 있습니다. 결정은 "무엇을 바꾸면 결과가 어떻게 되는가"에 답해야 내릴 수 있습니다. 앞의 질문은 상관으로 풀리고, 뒤의 질문은 상관으로 풀리지 않습니다.

이 구분을 건너뛴 채로 정확도만 올리면, 잘 맞히는데 아무도 따르지 않는 시스템이 남습니다.



상관은 알람을 울리고, 인과는 손잡이를 움직입니다.

1. 관측하는 질문과 개입하는 질문

같은 데이터를 놓고도 두 종류의 질문을 할 수 있습니다.

관측 질문: 온도가 높았던 날에 불량이 많았는가. 과거 데이터를 세면 답이 나옵니다. 상관계수 하나로도 됩니다.

개입 질문: 내일 온도를 2도 내리면 불량이 줄어드는가. 과거 데이터에는 이 답이 들어 있지 않습니다. 데이터에 기록된 것은 "온도가 그 값이었을 때 무슨 일이 있었나"이지 "누군가 온도를 그 값으로 바꿨을 때 무슨 일이 생겼나"가 아니기 때문입니다.

통계·인과추론 문헌에서는 이 둘을 관측 분포와 개입 분포로 나누고, 개입 쪽을 do 연산자 do(·)로 표기합니다. 이 표기를 바탕으로 개입 효과를 관측 데이터만으로 계산할 수 있는지 판별하는 규칙 체계가 do-calculus입니다. 표기법이 중요한 게 아니라, **두 질문이 같은 데이터에서 서로 다른 답을 낼 수 있다는 사실**이 중요합니다. 그리고 현장에서 실제로 필요한 것은 거의 항상 뒤쪽입니다. 사람이 알고 싶은 건 상태가 아니라 조치니까요.

한 문장으로 줄이면 이렇습니다. **상관은 지금 무슨 일이 일어나는지 알려주고, 인과는 무엇을 해야 하는지 알려줍니다.**

전에 대시보드가 늘어도 결정이 그대로인 이유를 쓰면서 화면이 늘어난 것과 판단이 나아진 것은 다르다는 이야기를 했습니다. 이 글은 그다음 칸입니다. 화면 대신 원인 설명을 붙여도 마찬가지라는 것, 즉 "왜 이렇게 됐는지"까지 설명해주는 모델도 여전히 관측 질문에만 답하고 있을 수 있다는 것입니다. 원인 설명과 개입 효과는 같은 말이 아닙니다.

2. 상관만으로 결정하면 무너지는 4가지

정확도를 아무리 올려도 메워지지 않는, 구조에서 오는 실패입니다.

(1) 뒤에 숨은 제3의 원인

여름이 되면 냉방 가동이 늘고 불량률도 오릅니다. 데이터만 보면 냉방과 불량이 함께 움직입니다. 여기서 "냉방을 줄이자"는 결정을 내리면 결과는 반대로 갑니다. 둘을 함께 밀어 올린 것은 습도였고, 냉방은 오히려 습도를 눌러주고 있었기 때문입니다. 이런 변수는 측정조차 안 되는 경우가 많아서, 데이터를 아무리 오래 모아도 자동으로 드러나지 않습니다.

(2) 방향이 반대인 관계

진동이 큰 설비일수록 정비를 자주 받습니다. 그 결과 데이터에는 "정비 이력이 많은 설비가 고장도 많다"가 남습니다. 이 상관을 그대로 읽으면 정비를 줄이자는 결론이 나옵니다. 실제 방향은 반대입니다. 고장이

잦아서 정비를 자주 한 것이지, 정비 때문에 고장이 난 게 아닙니다. 데이터는 이 방향을 스스로 알려주지 않습니다.

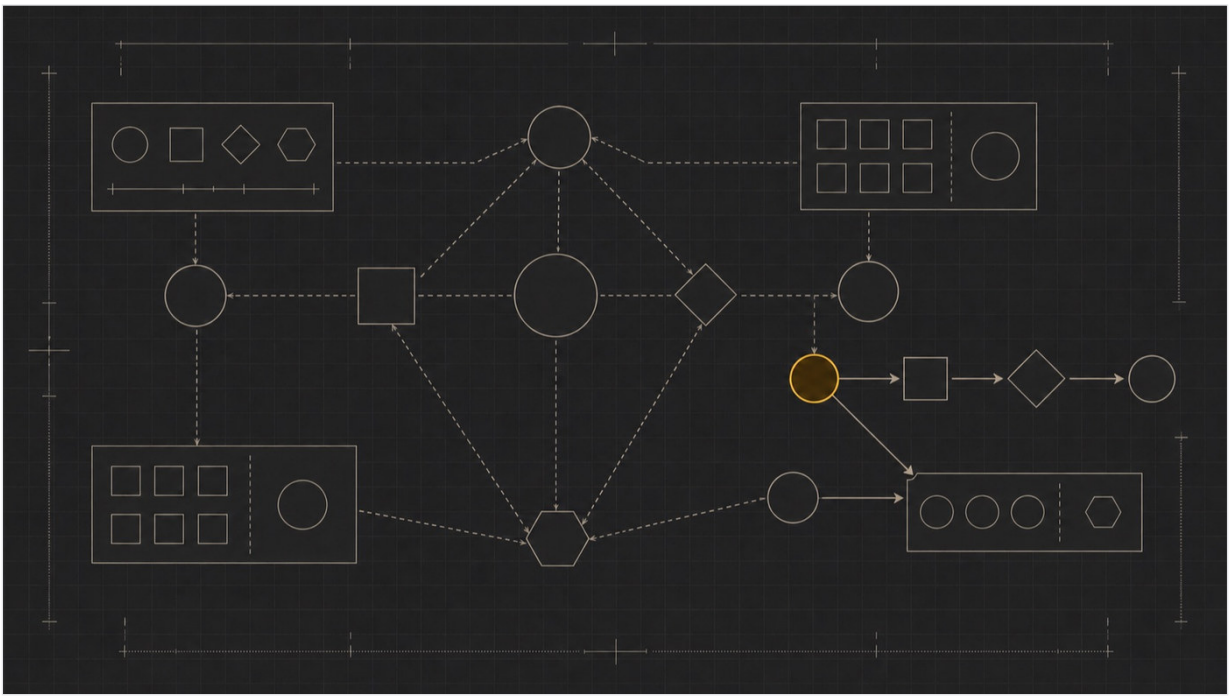
(3) 이미 걸러진 데이터

품질 검사를 통과한 제품만 후공정 데이터가 남는 라인이 많습니다. 그러면 앞 공정 조건과 최종 품질 사이의 관계가 통과분 안에서만 계산됩니다. 걸러진 집단 안에서는 원래 없던 관계가 생기기도 하고, 있던 관계가 사라지기도 합니다. 모델은 자기가 보는 데이터가 이미 선별된 것인지 알 방법이 없습니다.

(4) 개입하는 순간 관계가 바뀐다

모델이 학습한 것은 "그 조건에서 사람들이 그렇게 운전하던 시절의 패턴"입니다. 그 패턴을 근거로 설정을 바꾸면, 그 순간 데이터를 만들어낸 조건 자체가 달라집니다. 예측이 맞았는지 확인하려고 보면 이미 세상이 옮겨간 뒤입니다. 지표를 목표로 삼는 순간 그 지표가 지표 노릇을 그만두는 현상과 같은 자리에서 벌어집니다.

이 4가지의 공통점은 **데이터를 더 모아도 해결되지 않는다**는 것입니다. 부족한 건 표본이 아니라 구조 정보입니다.



관측하는 질문과 개입하는 질문은 그래프 위에서 다른 길을 갑니다.

---

3. 그래서 무엇을 더해야 하나

인과를 다루려면 데이터 밖에서 들어와야 하는 정보가 있습니다. 현실적으로는 3가지입니다.

(가) 무엇이 무엇에 영향을 주는지에 대한 그림. 설비·공정·품질 변수 사이의 영향 방향을 명시한 그래프입니다. 이걸 데이터로부터 완전히 학습해내는 방법은 아직 실무에서 믿고 사용할 수준이 아닙니다. 대부분은 현장 엔지니어의 지식을 받아 적는 데서 시작합니다. 이 단계에서 필요한 것이 설비·공정 용어를 하나로 맞춰두는 일이고, 그래서 온톨로지와 지식그래프 이야기가 인과 이전 단계로 먼저 나옵니다. 같은 대상을 라인마다 다르게 부르면 영향 관계를 그릴 대상 자체가 정해지지 않습니다.

(나) 개입을 실제로 해보는 방법. 인과를 확인하는 가장 확실한 방법은 바꿔보는 것입니다. 문제는 돌아가는 라인에서 조건을 임의로 흔드는 실험을 허락받기 어렵다는 점입니다. 그래서 개입 실험을 먼저 받아내는 자리가 디지털트윈이 됩니다. 다만 이때의 트윈은 지난 일정을 재생하는 화면이 아니라 조건을 흔들었을 때 응답이 달라지는 모델이어야 합니다. 이 구분은 트윈이 답을 미리 알고 있으면 시험이 아니다에서 따로 정리했습니다. 재생만 하는 트윈에 What-If 질문을 던지면 언제나 원래 답이 나옵니다.

(다) 개입 결과를 되받는 경로. 조치를 실행한 뒤 결과가 같은 체계로 돌아와야 다음 판단이 나아집니다. 알람만 내보내고 조치 결과를 받지 않는 시스템은 1년이 지나도 첫날과 같은 수준입니다(앞서 링크한 글에서 이 경로 이야기를 따로 다뤘습니다).

---

4. 해보고 나서야 알게 된 3가지

첫째, 상관 상위 변수만 모아 만든 모델은 계절이 바뀔 때 무너졌습니다. 검증 성적은 좋았는데, 실제로는 그 기간에만 함께 움직이던 변수들을 모아둔 상태였습니다. 성적표가 나빠서 알아챈 게 아니라 현장에서 "이 알람 왜 뜨는지 모르겠다"는 말이 반복돼서 알아챘습니다.

둘째, 원인 후보를 현장에 물어보면 데이터에 없는 변수가 먼저 나옵니다. 원자재 로트가 바뀌었다, 그날 작업자가 달랐다, 옆 라인 정비 때문에 압축공기 압력이 떨어졌다. 이 변수들은 어느 시스템에도 기록돼 있지 않았습니다. 인과 그림을 그리는 작업의 실제 산출물은 모델이 아니라 **"지금 우리가 측정하지 않고 있는 것들의 목록"** 이었습니다. 이게 나오면 계측 계획이 바뀝니다.

셋째, 인과라고 이름 붙인 결과가 더 위험할 수 있습니다. 상관 그래프는 보는 사람이 알아서 의심합니다. 그런데 화살표에 방향이 그려져 있으면 검토자가 덜 따집니다. 그 방향이 데이터에서 나온 것인지 누군가의 가정에서 나온 것인지 구분해 표기하지 않으면, 가정이 근거로 승격돼 돌아다닙니다. 그래서 영향 관계 하나하나에 근거 등급을 붙여두는 편이 낫습니다. 실험으로 확인된 것, 문헌으로 뒷받침되는 것, 현장 경험 기반의 가정, 3가지는 다르게 취급해야 합니다.

5. 판정 6문항 — 우리 AI는 알람인가 결정인가

지금 도입했거나 검토 중인 시스템에 그대로 물어보시면 됩니다.

1. 출력이 "이 값이 이상하다"에서 끝나는가, "이걸 이만큼 바꾸라"까지 가는가
2. 조치를 권할 때 그 조치를 실행했을 때의 예상 결과를 함께 내놓는가
3. 그 예상이 과거 유사 상황의 평균인가, 조건을 바꿔 돌려본 결과인가
4. 변수 사이의 영향 방향이 어딘가에 명시돼 있는가, 아니면 모델 안에만 있는가
5. 측정되지 않는 요인이 무엇인지 목록으로 관리되는가
6. 실행한 조치의 결과가 시스템으로 되돌아오는 경로가 있는가

~~12번만 되면 알람 수준입니다.~~ 3~~4~~번까지 가야 What-If 질문에 답하는 상태이고, 5~6번이 있어야 그 답이 시간이 지나도 틀어지지 않습니다. **6문항 중 4개 이상이 "아니오"인데 자율 판단을 맡기는 계획이 있다면, 순서가 뒤집힌 것입니다.**

6. 실무 순서

① **지금 나가는 알람이 어떤 결정으로 이어졌는지 30건만 손으로 추적합니다.** 알람 대비 조치 비율이 낮다면 문제는 정확도가 아니라 알람이 결정에 필요한 정보를 담고 있지 않은 것입니다.

② **가장 비싼 결정 1가지를 고르고, 그 결정에만 영향 관계 그림을 그립니다.** 전 공정을 한 번에 그리려 들면 끝나지 않습니다. 시작은 현장 인터뷰이고, 산출물은 그림과 함께 "측정하지 않고 있는 것들의 목록"입니다.

③ **그 그림을 조건 변경 시뮬레이션으로 검증합니다.** 응답이 현장 상식과 어긋나면 그림이 틀렸거나 상식이 틀린 것이고, 어느 쪽이든 알아낼 값이 있습니다.

④ **확인된 관계부터 조치 권고에 넣고, 나머지는 가정 표기를 유지합니다.** 전부 확인될 때까지 기다리면 아무것도 나가지 못합니다.

상관을 잘 찾는 일은 이제 어렵지 않습니다. 어려운 것은 그다음, 바꿨을 때 무슨 일이 생기는지에 답하는 일입니다. AI에게 판단을 맡기려는 계획이 있다면 정확도를 몇 퍼센트 더 올릴지보다 이 질문에 답할 수 있는 구조인지를 먼저 보는 편이 낫습니다. 알람은 사람을 부르지만, 결정은 사람을 대신합니다.

참고: Judea Pearl, "Causality: Models, Reasoning, and Inference" 및 do-calculus 관련 표준 문헌 · 교란·선택 편향·역인과 관련 통용 통계 문헌 — 2026-08-20 확인 기준. 본문의 공장 사례는 설명을 위한 예시이며 특정 고객사 데이터가 아닙니다.

제조 현장의 AI 적용 기록은 wace.me

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

원문 blog.vexplor.com/post/correlation-alerts-causation-decisions
발행 2026년 8월 19일 · 방동걸 (주)웨이스 대표이사

회사 주식회사 웨이스 (WACE Inc.) · 대표이사 방동걸 · 사업자등록번호 439-87-02852
주소 (18462) 경기 화성시 동탄구 동탄영천로 150, 에이동(A동) 1702-1708호 (영천동, 실리콘앨리동탄)
연락처 전화 070-4730-0629 · contact@wace.me
홈페이지 wace.me · linkedin.com/company/wace-inc

원문 홈페이지 링크드인