인사이트

SAP는 50년 노하우를 ERP의 경쟁력으로 내세우지만, 웨이스는 그 노하우의 값이 검증과 공장의 기록에서 나온다고 봅니다

SAP는 AI가 작은 애플리케이션은 복제해도 ERP는 쉽게 넘지 못한다고 말합니다. 그랑피에르 SAP 비즈니스 AI 제품 관리 총괄은 지난달 ZDNet Korea 인터뷰에서 기업이 AI로 ERP를 직접 만드는 문제를 두고 50년 동안 고객과 함께 쌓은 업무 프로세스와 애플리케이션...

방동걸 (주)웨이스 대표이사
2026년 11월 4일

[2026-11-05 추기] 이 글은 전날 확정한 원고를 예약대로 7시 1분에 공개했습니다. 같은 날 새벽 내부 실측에서 아래 시험이 잡지 못한 결함이 보고됐고, 공개 전에 원고에 반영하지 못했습니다. 화면이 부르는 API 길로 출하에서 매출전표를 이어 만들면 수량과 공급가가 0인 채 확정됩니다(내수도 같습니다. 원인은 코드를 읽어 세운 추정이고 브라우저 화면은 아직 보지 않았습니다). 4,210달러 수주 한 건은 환산 없이 4,210원으로 전표와 채권에 기표됐습니다(그 환율의 옳은 값은 5,705,476원). 확정한 매출전표의 통화와 환율이 수정되는 것도 확인했습니다. 앞의 두 건은 판정 프로그램의 식으로 계산해도 통과가 나왔습니다. 실제 운영 서버에 4,210달러 건과 같은 모양의 채권은 없었습니다. 같은 날 16시 조회에서 공급가나 줄 수량이 0인 확정 매출전표도 0건이었습니다. 다만 운영에는 출하에서 이어 만든 전표가 아직 없으므로, 이 값은 그 길을 지난 전표가 없다는 뜻이고 결함이 없다는 뜻은 아닙니다. 이 결함은 고치는 대상으로 남아 있습니다.

SAP는 AI가 작은 애플리케이션은 복제해도 ERP는 쉽게 넘지 못한다고 말합니다. 그랑피에르 SAP 비즈니스 AI 제품 관리 총괄은 지난달 ZDNet Korea 인터뷰에서 기업이 AI로 ERP를 직접 만드는 문제를 두고 50년 동안 고객과 함께 쌓은 업무 프로세스와 애플리케이션 로직을 근거로 들며 「이를 처음부터 다시 만드는 것은 그만한 가치가 없다」고 했습니다.

제조 AI 운영체제 VEXPLOR를 만드는 웨이스는 이 말에 절반은 동의하고 나머지 절반은 다르게 봅니다. 웨이스는 올해 8월부터 자사 제품 VEXPLOR Logic Studio의 솔루션 정의는 옮기고 실행 코드를 새로 작성했고, 10월에는 가상 회사의 1년을 13회 넘게 실행해 결과를 판정했습니다. 해 보니 코드를 작성하는 일보다 적힌 규칙이 맞는지 가려내는 일이 훨씬 어려웠습니다. 그래서 웨이스는 ERP의 경쟁 우위를 쌓인 햇수보다 검증과 공장의 기록에서 찾습니다. 다만 이것은 벤더가 자기 제품을 다시 작성한 사례이고 고객사의 자체 구축과는 다릅니다. 시험한 것도 웨이스의 제품이고, SAP 제품은 시험하지 않았습니다.

01

SAP는 50년 노하우를 ERP의 경쟁력으로 내세웁니다

SAP의 주장을 정리하면 세 가지입니다.

- **규모와 로직** — 대담 요약에 따르면 크리스티안 클라인 CEO는 ERP의 데이터 필드가 약 700만 개라고 말했고, 그래서 SAP 제품팀은 소프트웨어 거의 전부를 프롬프트로 다시 만들 수 없다고 결론 냈습니다.
- **검증과 운영** — ZDNet Korea 기사는 세금 계산처럼 항상 같은 결과가 필요한 업무는 규칙을 정확히 구현하고 검증해야 한다고 전합니다. 그랑피에르 총괄은 지속적인 운영 역량도 차이로 꼽았습니다.
- **정확도** — 클라인 CEO는 9월 Big Technology 대담에서 재무 마감 에이전트를 예로 들며 AI가 업무를 90% 이해하고 정확도는 93% 정도일 것이라고 어림했고, 100%가 필요하다고 말했습니다. 요약에 따르면 그는 몇 달 안에 도달한다고 봅니다.

세 가지 모두 일리가 있습니다. SAP도 구현이 쉬워졌다는 점은 인정하고, 규칙을 검증하고 운영하는 일은 대체하기 어렵다고 봅니다. 웨이스는 그 검증을 실행하는 비용도 내려가는 방향이라고 봅니다. 웨이스가 직접 해 본 범위에서는 비용의 구성이 달라졌다고 봅니다. 사람이 같은 시험을 실행한 비교값은 없습니다.

02

웨이스는 솔루션 정의를 옮기고 실행 코드를 새로 작성했습니다

Logic Studio의 옛 코드는 약 124.7만 줄이었습니다. 21개 솔루션의 테이블과 관계 등을 적은 정의 코드가 그중 229,712줄입니다. 웨이스는 이 정의를 데이터로 뽑아 옮기고, 그 정의를 읽어 실행하는 런타임과 화면, 업무 규칙은 새로 작성해 시험했습니다.

구분	옛 코드	재작성
전체 코드	약 124.7만 줄	약 13만 5천 줄
솔루션 정의	코드 229,712줄	데이터로 옮김
런타임 첫 계획(8월 22일)	해당 없음	1~2만 줄
런타임 지금(화면 제외)	해당 없음	약 9만 8천 줄

두 줄 수를 그대로 비교할 수는 없습니다. 옛 코드에는 정의 코드가 들어 있고, 재작성 수치는 11월 4일 기준으로 시험 시뮬레이터를 뺀 값입니다. 집계한 파일 종류도 옛 값은 ts·tsx·py·sql, 재작성 값은 ts·tsx·mts로 다릅니다. 이 표로 두 제품의 기능이 같다고 말할 수도 없습니다.

이 표에서 SAP의 말과 통하는 부분은 프롬프트 한 번으로 끝나지 않았다는 점입니다. 8월 22일 첫 계획은 런타임을 1~2만 줄로 잡았습니다. 그 계획은 일정과 인력 계산이 맞지 않아 같은 날 내부 검토에서 성립하지 않는다는 판정을 받았습니다. 지금 화면을 뺀 런타임 패키지만 약 9만 8천 줄로 그 상한의 약 5배입니다. 첫 계획의 배분에는 화면 렌더러도 들어 있어 층이 꼭 같지는 않습니다. 화면 등을 합친 제품 코드는 약 13만 5천 줄입니다. 범위가 넓어진 몫이 섞여 있어 어림이 그만큼 틀렸다는 뜻은 아니지만, 프롬프트 한 번으로 끝나는 일은 아니었습니다. 빨라진 부분도 있습니다. 재작성을 지시한 다음 날인 8월 23일에 사람이 로그인해 데이터를 만들고 고치고 지울 수 있었습니다. 다만 그날 된 것은 데이터와 화면까지였습니다. 지금 공개된 Logic Studio Cloud의 규모는 물리 테이블 1,539개, 화면 1,766개이고, 재작성한 제품은 아직 실제 공장에서 운영한 실적이 없습니다.



옛 코드에서 솔루션 정의는 데이터로 옮기고 실행 코드와 업무 규칙은 새로 작성한 구조

03

적혀 있던 규칙 35건을 그대로 켤 수는 없었습니다

옛 코드의 정의에는 솔루션을 넘나드는 업무 규칙 35건이 적혀 있었습니다. 어떤 표에서 어떤 사건이 일어나면 어느 표에 무엇을 할지, 필드 매핑과 조건까지 적혀 있어 완전해 보였습니다. 웨이스는 재작성 지시 사흘 뒤인 8월 25일에 이 35건을 분야별 AI 검토 세션으로 나눠 새 제품의 스키마, 시드, 실행 코드와 대조했습니다.

8월 25일 기준으로 새 제품에서 바로 켤 수 있는 규칙은 0건이었습니다. 이 수치는 그날의 판정이고 지금 제품 상태를 가리키지 않습니다. 막힌 이유 가운데 이 글에서 볼 네 가지는 이렇습니다.

- **실행 구조** — 20건은 상태 변경이나 수정 때 규칙을 실행할 장치가 그때 제품에 없었습니다.
- **기준 정보** — 같은 회사가 공급사 표와 거래처 표에 서로 다른 번호로 들어 있어 이름이 같아도 값이 통하지 않았습니다.
- **중복** — 컬럼이 다 맞았던 2건은 이미 다른 규칙이 하는 일이라, 켜면 재고가 두 번 잡히는 규칙이었습니다.
- **뜻** — 낙찰을 곧바로 발주로 만드는 규칙은 승인 없는 발주를 내보냅니다. 낙찰은 선정이고, 발주는 금전 채무를 지는 확약입니다.

이 글에서 따지려는 것은 마지막 두 가지입니다. 이름까지 맞아 보여도 켜면 안 되는 규칙이 있었고, 이런 규칙은 적힌 내용을 읽어서는 드러나지 않고 실물과 대조해야 드러났습니다. 실물과 대조해 확인하기 전의 규칙은 아직 지식이라고 부르기 어렵습니다. 웨이스가 노하우의 값이 햇수만으로 정해지지 않는다고 보는 첫 번째 이유입니다.

04

클라인 CEO가 어림한 93%는 AI가 답을 내는 정확도입니다

클라인 CEO가 어림한 93%는 재무 마감 에이전트를 예로 든 AI의 정확도입니다. 재무 마감에서 100%가 필요하다는 말에는 웨이스도 동의합니다. 클라인 CEO가 이 수치를 코드 작성에 적용한 것은 아니지만, AI 정확도 논의가 코드 작성까지 번질 때 이 수치를 그대로 옮길 수는 없습니다.

구분	에이전트가 매번 판단	코드로 작성해 실행
같은 입력	답이 달라질 수 있음	같은 답이 나옴
오류 모습	확률로 흩어짐	같은 곳에서 되풀이
찾는 방법	표본 평가	시험으로 재현
고치는 방법	모델·맥락 조정	코드 수정 후 재시험

오른쪽 열은 사람이 쓴 코드에도 똑같이 해당합니다. AI가 쓴 코드도 실행 단계에서는 오래된 검증 방법으로 다룰 수 있다는 뜻입니다. 물론 AI는 코드를 쓰고 고칠 때마다 새 오류를 넣을 수 있고, 시험 범위 밖의 오류는 재현되지 않습니다. 그래서 따져야 할 것은 AI의 정확도보다 그 규칙이 맞는지 누가 어떻게 확인하느냐입니다.

05

가상 회사 시험에서 실패가 0건이어도 「믿을 수 없음」 판정이 나왔습니다

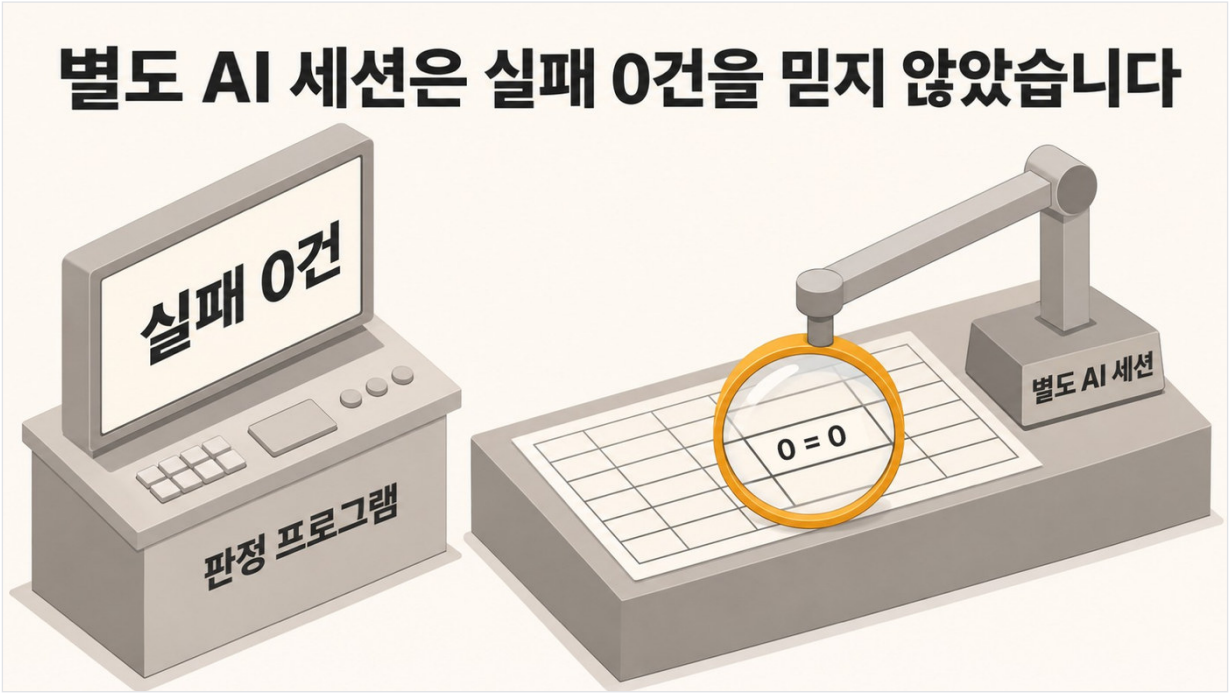
웨이스는 가상 제조사 하나를 만들어 재작성한 제품으로 1년을 운영하는 시험을 반복했습니다. 이 회사는 약 175명, 32개 부서 규모이고, 영업일 245일 동안 21개 솔루션 전부를 사용합니다. 업무를 실행하면 시험 판정 프로그램이 결과를 정답과 비교합니다. ~~510회차에는 통과 판정만 회차마다 약 83만~~84만 건 나왔습니다. 판정 프로그램이 틀린 값에 빨간불을 내는지도 25가지 사례로 함께 확인합니다. 정답 쪽 사본을 일부러 틀리게 만들거나 답을 아는 값을 넣는 방식입니다. 제품에 결함을 넣어 잡아내는지 본 시험은 아닙니다.

회차	판정 프로그램	독립 검증
1회차	실패 2,583건	해당 없음(결함 수정)
6회차	실패 0건	믿을 수 없음
12회차	판정 기준 충족	믿을 수 없음
10월 24일 2회	제품 원인 실패 0건	검증 항목 일치

표의 회차들은 전표를 직접 입력해 출하에서 매출전표를 이어 만드는 길을 지나가지 않았습니다. 「제품 원인 실패 0건」은 그 회차 판정 범위의 값이고 이 결함이 없다는 뜻이 아닙니다.

6회차는 실패가 0건이었지만, 판정에 관여하지 않은 별도 AI 세션은 믿을 수 없다고 결론 냈습니다. 생산오더 2,166행의 실적이 전부 0이었고, 정답도 같은 저장값을 읽어 0과 0이 같다는 통과가 1,433건 나왔기 때문입니다.

웨이스는 이 뒤 시험을 끝내는 기준을 바꿨습니다. 한 판정 항목의 통과가 전부 0과 0의 비교이면 통과로 치지 않아야 하고, 판정 프로그램이 확인하지 못한 항목은 원인을 모두 분류해야 합니다. 판정에 관여하지 않은 세션이 숨은 통과를 찾지 못해야 하고, 다른 조건의 회차에서도 같은 결과가 나와야 합니다. 다만 항목 안에 0이 아닌 건이 섞이면 0과 0인 건은 통과로 남습니다. 건별 기준은 아직 없습니다. 12회차는 판정 기준을 모두 충족했지만 이 독립 검증에서 숨은 결함 15건이 나왔습니다. 납품서 526건의 품목과 금액이 0이었는데, 판정 프로그램은 그 표를 보지 않았습니다.



판정 프로그램은 실패 0건을 냈지만 독립 검증이 0과 0의 통과를 찾아낸 시험 구조

납품서 결함을 포함한 제품 쪽 결함 10건도 고쳤습니다. 표에서 10월 24일 한 회차의 값은 실패 5건과 끊김 1건을 시험 환경의 데이터베이스와 시뮬레이터 원인으로 분류한 뒤의 결과이고, 이 회차 뒤에도 미충족 분류와 제품 보완은 계속됐습니다.

이 시험으로 알게 된 것은 두 가지입니다. 1년치 업무를 짧은 기간에 13회 넘게 실행할 수 있었고, 웨이스는 시험을 실행하는 비용이 크게 내려갔다고 봅니다. 사람이 같은 시험을 실행한 비교값은 없습니다. 하지만 무엇을 믿을지 정하는 기준은 회차를 거듭하고 나서야 드러났습니다. AI가 실행은 빨리 해 주지만, 판정 기준을 세우는 일은 여전히 어렵습니다.

이 시험에는 한계도 있습니다. 데이터는 모두 가상이고, 독립 검증도 AI 세션이었습니다. 회계·감사 전문가와 현장 담당자의 확인, 세무 신고, 데이터 이관, 장애 책임은 이 시험으로 확인하지 못했습니다.

06

SAP가 강한 부분은 이 시험이 대신하지 못합니다

SAP를 버리자는 이야기는 아닙니다. 다음 네 가지는 웨이스도 SAP와 의견이 같습니다.

- **큰 시스템은 코드 생성만으로 운영되지 않습니다** — 위 시험에서도 코드 생성만으로 끝나지 않았습니다.
- **운영 중인 ERP를 다시 만드는 일은 대부분의 회사에 맞지 않습니다** — 장애 위험이 큽니다. 웨이스는 타사 ERP를 복제하라고 권하지 않습니다.
- **법규 개정 대응과 감사 증빙, 장애 책임은 오래 운영한 벤더의 강점입니다** — 웨이스의 가상 회사 시험에는 이 경험이 들어 있지 않습니다.

- **규모의 차이는 큽니다** — 테이블 1,539개 제품의 실측만으로 700만 개 필드를 다루는 시스템을 단정할 수 없습니다.

기존 애플리케이션 위에 에이전트를 얹는 쪽에 새 가치가 생긴다는 SAP의 말에도 동의합니다. VEXPLOR도 없으면 구축하고, 쓰던 솔루션은 구축형(Enterprise)에서 교체 없이 연결합니다.

07

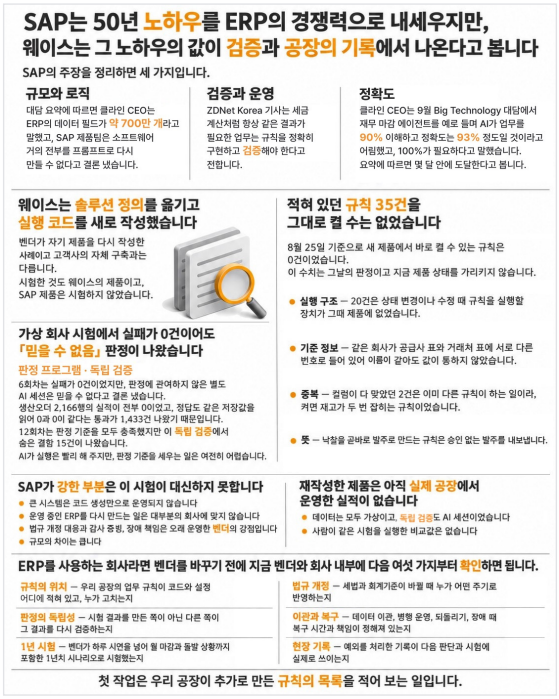
웨이스는 다음 경쟁이 공장의 기록을 시험으로 바꾸는 쪽에서 벌어진다고 봅니다

SAP의 50년에는 고객이 실제로 운영하며 확인한 결과가 들어 있고, 그 값은 큽니다. 다만 그 값은 햇수 자체보다 그동안 규칙을 실행해 확인한 데서 나옵니다. 확인 비용이 내려가면 햇수의 격차도 전처럼 고정되지는 않습니다. 또 표준 밖의 예외는 벤더의 코드보다 각 공장의 기록에서 나옵니다. 이 부분은 아직 웨이스의 판단이고, 실제 공장의 실측으로 확인한 결과는 없습니다.

웨이스도 이 방향으로 갑니다. 재작성 제품의 검증은 지금 가상 회사 시험까지 왔고, 다음 근거는 실제 공장에서 나와야 합니다. 웨이스는 현장의 예외와 판단 기록을 시험으로 바꿔 다음 판단에 사용하는 구조부터 만들려고 합니다.

ERP를 사용하는 회사라면 벤더를 바꾸기 전에 지금 벤더와 회사 내부에 다음 여섯 가지부터 확인하면 됩니다. 첫 작업은 우리 공장이 추가로 만든 규칙의 목록을 적어 보는 일입니다.

- **규칙의 위치** — 우리 공장의 업무 규칙이 코드와 설정 어디에 적혀 있고, 누가 고치는지
- **판정의 독립성** — 시험 결과를 만든 쪽이 아닌 다른 쪽이 그 결과를 다시 검증하는지
- **1년 시험** — 벤더가 하루 시연을 넘어 월 마감과 돌발 상황까지 포함한 1년치 시나리오로 시험했는지
- **법규 개정** — 세법과 회계기준이 바뀔 때 누가 어떤 주기로 반영하는지
- **이관과 복구** — 데이터 이관, 병행 운영, 되돌리기, 장애 때 복구 시간과 책임이 정해져 있는지
- **현장 기록** — 예외를 처리한 기록이 다음 판단과 시험에 실제로 쓰이는지



이 글을 한 장으로 요약하면 이렇습니다

웨이스가 보기에 ERP 노하우의 값은 쌓인 햇수만으로 정해지지 않고, 그 노하우를 실행해 확인하는 체계와 공장에 남는 기록에서 나옵니다. SAP의 50년도 그 확인을 거친 만큼 값을 합니다.

웨이스가 하는 일이 궁금하시면: wace.me

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

08

함께 읽어보세요

- AI가 코딩을 빠르게 해도, 제조 솔루션 구축 기간은 왜 그만큼 줄지 않을까
- ERP·MES·WMS가 서로 연결되지 않게 된 길을, 제조 AI는 왜 다시 밟고 있는가
- 직원이 편한 시스템을 만드는 중소·중견 제조기업은 담당자가 바뀔 때마다 처음으로 돌아갑니다

원문 blog.vexplor.com/post/sap-erp-know-how-verification-factory-records
발행 2026년 11월 4일 · 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