

AI 1호 직원의 5대 통제 - 실제 운영사례

AI의 능력과 권한을 분리하는 회사 내부통제 | 실제 사건 중심 요약

1 권한 통제

실행권은 자동 부여하지 않는다

실제 사건

'검토' 요청에서는 파일 수정·단계 이동을 금지하고, 외부 제출은 AI가 준비해도 대표 승인 전에는 실행하지 않았습니다.

대표 통제

검토와 실행을 별도 권한으로 나누고, 중요 실행은 명시적 승인 이후에만 허용했습니다.

통제 결과

AI의 수행능력과 실제 실행권을 분리해 사용자 의도보다 앞서가는 행동을 차단했습니다.

2 검수·복구 통제

오류를 최종업무로 통과시키지 않는다

실제 사건

공모전 마감정보를 23:59로 알고 있었으나 공식 접수화면 재검증으로 18:00을 확인했고, 파일·MVP 오류도 실제로 반려했습니다.

대표 통제

공식원문 재확인, 결과물 검수, 오류 반려퇴, 마지막 정상 기준점 복구 후 재작업을 적용했습니다.

통제 결과

잘못된 정보·파일·개발결과가 후속 업무의 기준으로 승계되는 것을 막았습니다.

3 기능·시스템 권한 통제

기능 보유와 사용 권한을 분리한다

실제 사건

AI가 편의를 위해 메모리를 변경하거나 신규 프로젝트를 앞서 생성하려는 상황을 권한 문제로 판단했습니다.

대표 통제

명시적 요청 없는 메모리 생성·수정·삭제를 금지하고, 신규 프로젝트는 운영관리 승인 게이트를 거치도록 했습니다.

통제 결과

AI가 가진 기능을 대표 승인 없이 임의로 사용하는 것을 방지했습니다.

4 정보 통제

내부 활용정보와 외부 공개정보를 분리한다

실제 사건

AI는 업무상 개인정보·사업기밀·내부 프롬프트·상세 데이터 구조를 참고할 수 있지만 외부 제출물에는 그대로 공개하지 않습니다.

대표 통제

공개 필요성을 별도로 판단하고 익명화·일반화·마스킹·부분 공개를 적용했습니다.

통제 결과

AI 활용 범위를 유지하면서도 내부 보안과 개인정보 노출 위험을 낮췄습니다.

5 윤리·책임 통제

사업성이 있어도 책임 위험이 크면 멈춘다

실제 사건

신사업 탐색에서 유가족·장례처럼 취약한 이용자와 관련된 민감영역은 개인정보·금융·상속, 분쟁 가능성을 별도 위험으로 보았습니다. 또한 대표의 비전문 영역은 고객 서비스의 책임경영에 위배되기에 GPT 답변만 믿고 신규사업을 개발·확정하지 않습니다.

대표 통제

시장성 외에 윤리·책임 리스크를 독립적으로 검토하고, 전문판단이 필요한 영역은 전문가 검증 없이 사업화하지 않는 기준을 적용했습니다.

통제 결과

"AI가 가능하다고 말한다"를 사업화 근거로 삼지 않고 사람의 책임판단과 전문성 한계를 최종 게이트로 유지했습니다.

AI에게 업무는 위임하지만 경영권은 위임하지 않는다.