증빙 3. AI 얼굴 모자이크 — 오류율을 재고 사람 검수 게이트를 만든 기록

2026 AI 활용 사례 공모전 출품작 1(업무 생산성) 첨부 자료 · 개인정보 보호와 책임 있는 AI 활용 사례

1. 경위

개인 홈페이지에 올리는 맛집·여행 사진에는 다른 손님과 직원들의 얼굴이 찍혀 있습니다. 초상권과 개인정보 보호를 위해 AI로 얼굴을 자동 검출해 모자이크했지만, 결과를 전수 육안으로 대조해 보니 AI가 얼굴이 아닌 것까지 대량으로 얼굴로 잘못 찾는다는 사실이 드러났습니다.

2. 측정 결과

항목	수치	설명
AI 검출 전체	513곳	자동 검출이 얼굴이라고 표시한 전체
오탐(얼굴 아님)	218곳 (42%)	음식·물건·캐릭터 등을 얼굴로 오인
오탐 사진 복원	134장	얼굴이 없는데 모자이크된 사진을 원본으로 복원
혼합 사진 재처리	26장	실제 얼굴과 오탐이 섞인 사진은 실제 얼굴만 다시 처리
영향 글 반영	57개 글	수정 결과를 글 전체에 반영하고 재검증

3. 설계한 게이트 (기준은 낮추지 않고, 적용 전에 사람이 막는 구조)

검출 기준을 올리면 진짜 얼굴을 놓쳐 개인정보 사고로 이어지므로 기준은 그대로 두었습니다. 대신 적용 절차를 바꿨습니다. ① AI가 검출하면 자동 보류 상태가 됩니다. ② 사람이 전수 검수해 승인한 파일만 적용 대상이 됩니다. ③ 승인 파일이 없으면 적용 프로그램 자체가 실행을 중단하도록 강제했습니다. 이 구조 덕분에 오탐 42%라는 AI의 한계를 안고도, 개인정보 보호와 사진 품질을 둘 다 지킬 수 있었습니다.

4. 원칙

AI의 결과는 '확인했어요'라는 말로 통과시키지 않습니다. 오류율을 수치로 재고, 틀릴 수 있음을 전제로 사람이 확인하는 자리를 절차 안에 만듭니다. 이 원칙은 모자이크뿐 아니라 발행 검수·성과 측정 등 운영 전반에 동일하게 적용하고 있습니다.