CMDSPACE LABS | 공개 보고서

국내외 금융·보험 AX 사례와 보안 대응

2026-10-07

구요한 (CMDSPACE)

웹으로 읽기: https://labs.cmdspace.work/kb-insurance-ax-1007/

직접 출처 27건(S01~S27)은 10절에 주소와 함께 실었습니다.

1. 임원이 먼저 읽을 핵심 판단

금융·보험의 AX는 AI를 업무 흐름에 연결하고, 그 결과를 누가 검토하며 어떤 권한으로 실행할지 정하는 과정이다. 공개 사례에서 반복되는 변화는 지식 검색, 보고서와 회의의 후속 업무, 상담원 지원, 제한된 처리 단계의 자동화다. 보험에 적용할 때는 업무의 속도와 함께 판단 근거·고객 영향·승인 책임을 설계해야 한다. 이 보고서는 국내 5개·해외 7개 사례를 통해 세 가지 경영 판단을 제안한다.

1. **근거를 확인할 수 있는 업무부터 시작한다.** 보고서 검토, 승인 지침 검색, 상담 후속 메모는 입력·출력·사람 검토를 구체적으로 정하기 쉽다.
2. **지식과 권한의 책임자를 함께 지명한다.** 현업 전문가가 판단 기준을 정리하고, IT·보안·준법 담당이 허용되는 자료와 실행 범위를 확인한다.
3. **확대는 검증 결과로 결정한다.** 초안 작성 속도, 사람의 검토·수정 시간, 중요한 오류, 유지 공수, 보안 위반을 함께 비교한다. 여기서 제안하는 적용 순서와 90일 계획은 실행 전 설계안이다. 기업 발표에 있는 기능·성과와 본 보고서의 해석을 구분해 읽어야 한다. 회사의 높은 도입률이나 시간 절감 수치를 KB손해보험의 예상 ROI로 직접 옮길 근거는 없다.

2. 금융 업무에서 AX가 바꾸는 것

세 가지 활용 수준

AI의 활용 수준에 따라 검증 대상과 승인 구조가 달라진다.

활용 수준	바뀌는 업무	보험 업무의 예	관리할 핵심
분류·예측 AI	위험을 점수화하고 이상을 찾는다	사기 위험 탐지, 손해 위험 분석	표본 편향, 오탐·미탐, 모델 변화
생성형 AI 지원	자료를 찾아 요약하고 초안을 만든다	지급 지침 검색, 보고서 검토, 상담 요약	근거, 문서 버전, 접근권한, 사람 검토
에이전트 실행	여러 도구를 호출해 업무 단계를 진행한다	후속 업무 등록, 제한된 청구 처리 지원	도구별 권한, 실행 승인, 중복 방지, 복구

위 구분은 사례를 읽기 위한 분석 틀이다. 제품에 'Agent'라는 이름이 붙었다는 사실만으로 자율 실행 범위를 알 수는 없다. 실제로는 읽기, 초안 작성, 등록, 발송, 금액 결정, 지급 각각의 권한을 확인해야 한다.

보험 가치사슬에 적용하는 순서

다음은 공개 사례에서 도출한 적용 제안이다. 각 업무의 현행 승인과 자료 처리 조건을 먼저 확인한다.

업무	시작할 수 있는 지원 과제	현업이 제공할 지식	별도로 승인할 실행
상품·영업	상품 조건 비교, 설명 초안	승인된 약관·설명자료·금지 표현	고객에게 보내는 설명·권유
인수심사	자료 정리, 근거 탐색, 검토 의견 초안	위험 판단 기준·예외·추가 확인 조건	인수·거절·조건 변경
고객상담	지침 검색, 상담 요약	시행 중인 업무 지침·응대 기준	고객 통지·계약 정보 수정
보험금 처리	청구 자료 정리, 지급 지침 조회	보장·면책·금액 산정 규칙	지급·부지급·감액 결정
경영관리	보고서 검토, 회의 질문, 후속 메모	지표 정의·판단 기준·책임 체계	대외 보고·경영 지시의 확정
IT·보안	로그 분석, 취약점 우선순위 제안	자산 목록·대응 기준·복구 절차	계정 잠금·격리·패치 배포

임원에게 필요한 첫 질문은 “어느 AI를 구매할 것인가?”와 함께 “어느 업무의 어떤 판단을 지원하고, 누가 결과를 책임질 것인가?”다. 이 답이 있어야 자료 정리·제품 선택·성과 평가가 같은 방향으로 진행된다.

3. 국내 사례: 현업 지식과 실행 통제가 만나는 지점

3.1 KB손해보험 — 에이전트의 직무·권한·이력을 관리한다

KB손해보험은 2026년 9월 공개 자료에서 AI 에이전트 6종에 사원번호를 부여하고 업무에 적용한다고 설명했다. 공개된 업무에는 과실도표 검색, 상담·민원 요약과 분류, 후속 등록 지원, 사고접수 정리, 지급 지침 검색 등이 포함된다. 직무와 권한, 업무 이력을 관리한다는 점이 임원 교육에 유용하다.[S01] 회사는 ‘직원 1명의 한 달 치 업무 이상’이라는 평가도 제시했지만, 산식·표본·검토와 운영에 드는 시간을 포함한 순절감은 공개하지 않았다. 에이전트 평가·등급제는 발표 당시 향후 계획으로 구분해야 한다. 경영진이 가져갈 질문은 구체적이다. 각 에이전트의 업무 소유자는 누구인가? 어떤 자료에 접근하며, 어떤 도구를 실행하는가? 오류가 발견되면 누가 권한을 회수하고 업무를 복구하는가? 사원번호와 함께 이 책임 구조가 있어야 운영 수준을 판단할 수 있다.

3.2 KB손해보험 — 현업 자율개발을 운영 가능한 과제로 연결한다

2026년 9월 ‘언더커버 에이전터스’ 행사에서는 현업이 M365 Copilot으로 업무 에이전트를 설계·구현하고, 임원이 미션 스폰서로 참여했다. 공식 발표에서 10개 팀의 참여와 생산성 분석 에이전트의 수상을 확인할 수 있다.[S02] 이 사례는 현업이 문제를 정의하고 전문가의 지식을 구체화하는 출발점을 보여준다. 행사에서 만든 결과가 전사 운영으로 이어지려면 검증, 자료 사용 승인, 운영 이관, 지식 갱신에 필요한 시간과 책임이 배정돼야 한다. 행사 결과 자체는 운영 효과 연구와 구분한다.

3.3 신한은행·LG CNS — 금융지식을 검색 가능한 업무 자산으로 만든다

LG CNS는 2025년 5월 신한은행에 금융지식 Q&A 서비스를 구축했다고 발표했다. 10만 건 이상의 업무·상품 지식을 대상으로 관련 문서를 찾아 답변에 사용하는 RAG와 개인정보·이상행위 차단 설계를 설명했다.[S03] RAG는 질문과 관련된 자료를 검색해 AI의 답변에 연결하는 방식이다. 보험에서도 지급 지침·업무 매뉴얼을 찾는 시간을 줄이는 후보가 될 수 있다. 다만 문서 수가 많아질수록 승인본, 시행일, 예외, 사용자별 접근권한을 관리해야 한다. 공급자의 보안 기능 설명과 실제 공격에 대한 방어율은 서로 다른 증거다.

3.4 신한은행 — 여신 전문가의 판단 준비를 지원한다

신한은행은 2026년 3월 생성형 AI 기반 여신심사지원 Agent 도입을 발표했다. 기업정보·재무·업종 자료를 정리하고 여신 의견서 작성을 지원하며, 12개 산업의 특화 분석 로직과 현업 노하우를 결합했다고 설명한다.[S04] 보험 인수심사에 연결할 수 있는 지점은 전문가가 검토할 자료와 판단 기준을 구조화하는 일이다. 공개 자료가 설명하는 범위는 분석과 의견서 지원이다. 자동 승인 범위나 승인 정확도의 개선은 별도 증거가 필요하다.

3.5 카카오뱅크 — 대화형 입력과 이체 확인을 연결한다

카카오뱅크의 2026년 8월 공식 안내는 대화형 이체 과정에서 수취정보를 확인하고 사기 의심·착오송금 경고를 제공하는 동선을 설명한다.[S05] 보험 업무에서 참고할 부분은 사용자의 의도를 이해하는 단계와 실제 돈을 보내는 단계를 나누는 설계다. 지급 대상, 계좌, 금액을 사용자가 확인하고 승인하는 과정이 필요하다. 해당 서비스의 현재 고객별 한도·인증 구현이나 사고 감소 효과까지 이 안내만으로 확정하지는 않는다.

4. 해외 사례: 지원에서 제한 실행까지

4.1 Morgan Stanley — 회의 이후의 업무를 바꾼다

Morgan Stanley는 2024년 6월 AI Assistant와 Debrief를 설명하면서 지식 검색, 고객 동의를 받은 회의의 메모, 후속 이메일 초안, CRM 기록을 연결했다. 담당 Financial Advisor는 이메일을 수정하고 발송 여부를 판단한다.[S06] 이 사례의 교육적 가치는 회의 기록을 다음 업무로 이어가는 흐름에 있다. 보험 임원에게는 보고서 검토 → 확인 질문 → 회의 메모 → 승인된 후속 조치로 연결해 설명할 수 있다. 각각의 산출물에 검토자가 있고, 외부 발송은 별도의 결정으로 둔다. 성과를 인용할 때는 분모를 붙인다. 발표의 **98%는 Financial Advisor 팀의 Assistant 채택률**이다. 회의당 30분 절약은 한 사용자의 경험 인용이며 전체 사용자의 평균 효과와 구분한다.

4.2 DBS — 여러 AI 과제와 운영모델을 함께 관리한다

DBS의 2025 연차보고서는 데이터 분석·AI/ML 포트폴리오에서 약 SGD 10억의 경제가치를 추정하고, 2,000개 이상의 모델과 430개 이상의 활용 사례를 설명한다. 이 수치는 여러 과제의 회사 추정값이며 생성형 AI 단독 ROI로 읽으면 안 된다.[S07] 구체적인 업무를 보면 구분이 더 선명하다. 2024 CIO statement의 CSO Assistant는 싱가포르·홍콩·인도·대만 상담원이 통화 내용을 실시간으로 문자화하고 요약·해결안을 추천받도록 지원한다. 상담 처리시간 **최대 20% 단축은 당시 기대값**이다.[S23] 같은 문서의 변경요청 100% 점검과 변경요청으로 인한 월평균 사고 수 81% 감소는 별도의 AI/ML 변경 위험평가 사례다. 비교 기간·표본·통제집단·산식이 공개 본문에 제시되지 않아 인과 효과를 확정하기 어렵다. 더 넓은 운영 재설계로 누적 130만 시간 이상을 절약했다는 회사 발표 역시 개별 상담 지원의 효과와 구분한다. DBS에서 배울 점은 과제별 효과를 정의하고, 조직·프로세스·현업의 학습과 함께 관리하는 방식이다. 서로 다른 업무의 수치를 한 AI 서비스의 성과로 합산하지 않아야 한다.

4.3 Allianz EKA — 상담원의 지식 탐색을 지원한다

Allianz는 2025년 2월 Enterprise Knowledge Assistant(EKA)를 소개하면서 2024년 독일 Leben 적용을 설명했다. 상담원이 관련 지식을 찾고 요약하며 원문 링크를 확인하는 구조다.[S09] 보험 지침 검색에 가까운 사례지만, 회사가 발표한 '일반 질문 정확도 90% 초과'는 시험 표본과 분모가 공개되지 않은 수치다. 이를 보험금 지급 판단의 정확도나 다른 회사에서 보장되는 성능으로 확대할 수 없다. 실제 적용에서는 문서의 최신성, 근거 일치, 권한 준수, 답변을 유보해야 할 질문을 함께 평가해야 한다.

4.4 Allianz Nemo — 처리 단계를 자동화하고 최종 지급은 사람이 결정한다

Allianz의 Project Nemo는 2025년 7월 호주에서 가동했다고 발표된 사례다. 대상은 자연재해 정전과 관련된 **AUD500 미만의 식품 손상 소액 청구**다. 7종 에이전트가 보장·날씨·사기·금액·감사 등을 지원하고, 최종 지급 결정은 사람이 한다.[S27] 회사는 처리·정산 시간 80% 감소를 발표했지만 표본·비교 기간·산식은 공개 본문에서 확인되지 않았다. '5분 미만'은 예시 청구가 최종 사람 검토로 전달되는 흐름의 설명이다. 이 사례는 자동화할 업무의 복잡도와 범위를 제한하고, 에이전트의 추천과 사람의 지급 승인을 나누는 설계를 보여준다. KB 적용 후보를 논의할 때도 청구 유형, 예외 조건, 금액과 고객 영향, 중단과 복구 경로를 먼저 정해야 한다.

4.5 AXA — 직원이 사용하는 업무 환경에 AI를 배포한다

AXA는 2026년 7월 Secure GPT 기반과 Microsoft 365 Copilot의 점진 배포, 교육, 중요한 판단에서의 인간 감독을 설명했다.[S08] Word·Teams·Outlook 같은 업무 환경에서 시작하는 접근은 익숙한 작업에 AI 지원을 붙이는 후보가 된다. 공개 발표는 점진 배포의 설명이다. 전 직원의 배포 완료나 업무별 효과를 검증한 결과와 구분하고, 도입 기업에서는 실제 허용 기능·계정·자료 범위를 확인해야 한다.

4.6 JPMorganChase — 공통 플랫폼 위에서 부서별 활용을 확장한다

JPMorganChase의 2024 연차보고서 관련 서한은 LLM Suite를 20만 명 이상의 직원에게 배포했다고 설명한다.[S10] 여기서 읽을 수 있는 것은 배포 규모와 통제된 공통 플랫폼의 접근이다. 실제 활성 사용자, 매일 사용하는 비율, 과제별 순절감은 별도 지표다. 보험사에 적용한다면 공통 계정·보안·기록 체계 위에 부서별 승인 지식과 평가 기준을 결합하는 구조를 검토할 수 있다.

4.7 Mastercard — 탐지 성과는 오탐과 손실까지 함께 본다

Mastercard의 2024년 2월 Decision Intelligence Pro 발표는 초기 모델링에서 사기 탐지율이 평균 20%, 일부 사례에서 300% 개선됐다고 설명했다. 발표 당시의 모델링 결과와 출시 예고라는 범위를 붙여 읽어야 한다.[S11] 보험사기 탐지에 연결할 때는 탐지율 하나만으로 성공을 판단하기 어렵다. 정상 고객을 잘못 차단하는 비율, 놓친 사기, 사람 재심의 부담, 최종 손실과 고객 피해를 함께 비교해야 한다. 거래 사기 모델링의 수치를 보험사기 예방률로 직접 환산할 근거는 없다.

5. 성과 수치를 읽는 기준

수치는 무엇을 측정했는지에 따라 의미가 달라진다. 발표 주체, 업무 범위, 분모, 비교 기준을 먼저 확인해야 한다.

공개 수치	정확히 읽을 범위	경영 판단에 추가로 필요한 정보
KB 에이전트 6종·업무량 평가	업무 적용과 회사 평가	실제 사용량, 검토·유지 공수, 오류, 순절감
Morgan Stanley 98%	Financial Advisor 팀의 Assistant 채택	활성 사용, 과제별 효과, 품질
DBS 약 SGD 10억	데이터 분석·AI/ML 포트폴리오의 추정 경제가치	과제별 산식·비용·중복 계산 여부
DBS 상담시간 최대 20%	CSO Assistant의 당시 기대값	실제 달성값, 비교 기간, 업무 난이도
DBS 사고 수 81% 감소	변경요청으로 인한 월평균 사고 수의 회사 설명	기간, 표본, 비교 설계, 다른 변화의 영향
Allianz EKA 90% 초과	일반 질문에 대한 회사 자체 정확도 설명	질문 표본, 정답 판정, 예외·권한 오류
Allianz Nemo 80% 감소	제한 청구 처리·정산 시간의 회사 발표	비교 기간·표본·산식·재작업
JPMorganChase 20만 명 이상	플랫폼 배포 대상 규모	활성 사용자, 실제 업무 활용, 순절감
Mastercard 평균 20%·일부 300%	초기 모델링의 탐지율 개선	운영 성과, 오탐, 손실, 고객 영향

출처는 각 사례의 각주에 연결했다. 이 자료에는 동일 조건의 독립 효과 비교나 신뢰구간이 제공되지 않았다. 따라서 도입률·배포 규모·기대값·초기 실험·회사 추정값을 같은 수준의 효과 증거로 취급하지 않는다. KB에서 검증할 질문은 "업무 한 건의 총시간이 줄면서 품질과 보안이 유지되는가?"다. **총시간에는 자료 준비, AI 실행·대기, 사람 검토, 수정·재작업, 후속 승인을 모두 포함**한다. 현업 지식 갱신과 운영·검증에 추가되는 공수도 별도로 기록한다. 시간 절감만으로 인력 감축이나 손해율 개선을 결론 내리지 않는다.

6. 보안 대응: AI 활용의 통제와 AI 공격에 대한 방어

6.1 AI가 읽는 자료부터 실제 실행까지 통제한다

AI의 안전성은 모델, 문서, 사용자 권한, 연결 도구, 승인, 기록, 복구의 전체 경로에서 평가해야 한다. 아래는 OWASP와 금융보안원 자료를 보험 업무에 연결한 설계 제안이다.[S18,S19,S15]

위험 장면	필요한 통제	확대 전에 요구할 증거
고객·건강·신용정보가 승인 밖 서비스로 전송됨	허용 계정·자료·목적·처리 경로, 최소 항목, 보관·삭제 조건	실제 데이터 흐름과 준법·개인정보 검토
사용자가 무관한 고객·임원 자료를 검색함	원문·검색 결과·캐시의 권한 관리, 공유 링크 정리	무권한 계정의 접근 차단 시험
첨부문서 속 지시가 업무 명령으로 실행됨	자료와 실행 지시의 분리, 도구 제한, 고위험 실행 승인	합성 악성문서의 유출·무승인 실행 차단 결과
요약 에이전트가 지급·발송까지 수행함	읽기·초안·등록·발송·지급 권한 분리, 대상·금액 확인	도구 허용 목록, 승인·철회 기록
구지침·미승인 초안이 현행 기준으로 답변됨	승인 상태·시행일·버전·근거 표시, 충돌 시 판단 유보	정상·예외·문서 충돌 검증 결과
잘못된 금액이나 계약 조건이 확정됨	검증 가능한 계산·업무 규칙과 담당자 확인	계산 경로, 전제, 예외와 재현 결과
모델·연결 도구가 바뀌어 처리 범위가 확대됨	변경 승인, 계약·권한 검토, 회귀검증, 복구 버전	변경 전후의 차이와 재검증 기록
반복 실행·장애로 업무가 중단됨	중복 방지, 재시도 한도, 긴급 정지, 수동 대체	장애·취소·복구 훈련 결과

프롬프트 인젝션은 AI가 읽은 자료 안의 공격 지시가 원래 업무 규칙을 흔드는 위험이다. 검색한 문서를 답변에 연결하는 RAG도 이 위험의 경로가 될 수 있다. 내부망이나 검색 기능만으로 해결됐다고 판단하기보다, 공격 자료를 넣어 권한·실행·유출 통제가 작동하는지 확인해야 한다.[S18] 금융보안원의 AI 보안 안내서는 개발·검증·운영의 전 생명주기, 별도 전문인력의 점검, 위험도에 따른 제3자 검증을 권고한다. 임원에게 필요한 산출물은 정상 동작 시연과 함께 공격·오류·중단 시나리오의 시험 결과, 모델·시스템 프롬프트·RAG 자료의 소유자와 변경 이력, 긴급 정지와 사람 개입의 증거다.[S15]

6.2 회사 Copilot의 계약 조건과 실제 자료 흐름을 확인한다

Microsoft 문서는 조직 데이터의 권한에 따른 접근과 프롬프트·응답·Graph 자료를 기반 모델 학습에 사용하지 않는 조건을 설명한다. 상호작용 기록의 저장과 web search·agents·외부 모델의 처리 조건은 별도로 확인해야 한다.[S17] 도입 기업은 회사의 승인 계정, 테넌트와 계약, 보관과 처리 지역, 연결 도구, 외부 전송 경로를 검토해야 한다. 기존 문서의 공유 범위가 지나치게 넓으면 AI가 그 권한을 따르더라도 업무상 불필요한 자료가 검색될 수 있다. 따라서 자료 공유와 접근권한 정비가 도입 과정에 포함돼야 한다.

6.3 딥페이크·위조 신원과 AI로 강화된 공격에 대응한다

FinCEN은 생성형 AI 미디어와 위조 신분증이 금융기관의 신원확인·인증을 우회하는 사기에 사용되는 정황을 경보로 알렸다. 이는 공격 유형의 근거이며 전체 사기의 발생률이나 특정 탐지기의 성능 자료와 구분한다.[S21] 보험 업무의 대응 후보는 독립 채널의 본인확인, 지급계좌 변경·고액 지급의 추가 확인, 행동·거래 신호 대조, 사람 재심과 조사 절차에 따른 증거 보존이다. 영상·음성 판별 결과를 거래 승인과 함께 교차 확인하는 구조가 필요하다. 사이버 방어에서는 외부 노출 자산, 공급망, 취약점 우선순위, 패치 지연, 침해 확산과 복구를 관리한다. 금융보안원의 프런티어 AI 대응 요령은 경영진 책임과 공동 대응·복원력, 방어 자동화를 함께 제시한다.[S16] 로그 분석과 조치 추천에서 출발해 계정 잠금·격리·패치 배포로 확장할 때는 영향과 복구 가능성에 맞춘 검증·승인 기준을 정해야 한다. 해당 요령은 비규제 안내 자료다. 문서의 적용 예시인 시간·배점·자동화 수준을 모든 금융사의 법정 기준으로 사용하지 않는다.

7. 규제와 해외 거버넌스 자료를 해석하는 방법

국내: 연결, 자료 처리, 자동 실행을 각각 확인한다

"외부 서비스에 연결할 수 있는가", "해당 자료를 처리해도 되는가", "그 업무를 자동 실행해도 되는가"는 서로 다른 검토다. 망분리 예외와 함께 개인정보·신용정보·외주·보관·소비자 보호·내부 승인 조건을 확인해야 한다. 금융위원회는 2026년 6월 발표에서 금융분야 AI 가이드라인의 6월 22일 시행과 자율규제 성격을 설명했다. 거버넌스·합법성·보조수단성·신뢰성·금융안정성·신의성실·보안성의 7원칙을 제시하며, 해당 법률의 의무는 별도로 구분한다.[S13] 기존 혁신금융서비스의 모델 변경도 보안 영향에 따라 절차가 달라진다. 새 모델로 바꿀 때는 입력·출력·구성·권한·자료 처리의 변화를 확인하고, 지정과 보안 평가의 조건을 검토해야 한다.'[S12] 2026년 9월 발표된 제2차 보안 목적 망분리 테스트는 선정 회사의 보안 취약점 탐지·개선, 대체 통제, 한시 비조치의 조건을 설명한 자료다.'[S14] 10월 6일 한국금융신문은 금융위 관계자를 인용해 7일 참여사 선정 일정 조정을 보도했다.'[S26] 본 보고서의 재조회 범위에서는 공식 새 일정·회사별 발급을 확인하지 않았다. 실제 사용 판단에는 해당 회사의 발급·목적·조건·유효기간을 확인해야 한다.

해외: 책임·문서·검증을 비교할 기준으로 읽는다

해외 자료는 관할과 문서의 성격을 붙여 활용한다.

자료	읽을 핵심	적용 범위의 경계
EIOPA AI governance Opinion	보험 AI의 데이터·기록·공정성·설명·보안·인간 감독	AI Act상 고위험·금지 AI는 제외; 새 의무를 신설하지 않는 기존 보험법규의 감독 해석이며 한국 직접 적용 규정과 구분
NAIC Model Bulletin	위험에 비례한 서면 AI 운영·관리 프로그램과 검사 문서	Model Bulletin은 model law·regulation과 구분; 미국 주별 현행 적용은 추가 확인
MAS MindForge Phase 2	Handbook 2종과 Implementation Examples 1종을 구분해 활용	보험·자산운용·Agentic AI까지 확장한 Toolkit 개요; 세부 PDF 전체는 미대조
NIST AI 600-1	생성형 AI 위험의 식별·평가·관리	자율적 참조 프로필; 인증이나 법적 적합성 보증과 구분

자료의 출처는 각각 EIOPA·NAIC·MAS·NIST다.[S22,S24,S25,S20] 책임 배정·검증·변경관리는 본 보고서가 제안하는 활용 질문이다. 이 자료들을 고객사에 연결할 때는 "우리 업무·지식·제3자 책임과 검증 결과를 문서로 설명할 수 있는가?"를 공통 질문으로 사용할 수 있다.

8. KB손해보험 임원 교육과 90일 실행에 연결하기

교육에서 가져갈 세 가지 장면

첫째, **Morgan Stanley의 회의 후속 업무**를 Word 보고서 검토와 연결한다. 합성 보고서에서 근거가 약한 주장, 빠진 확인 질문, 회의 후속 메모를 만들고 담당자가 검토하는 흐름을 보여줄 수 있다. 둘째, **DBS의 과제별 운영·성과관리**를 통해 도입 규모와 효과를 구분한다. 같은 'AI 성과'라는 이름 아래 상담 지원의 기대값, 변경관리 지표, 포트폴리오 경제가치가 섞일 수 있음을 설명한다. 셋째, **Allianz EKA와 Nemo의 지식·승인 구조**를 보험 업무에 연결한다. 지급 지침의 승인본과 미승인 개정 초안을 구분하고, AI의 검증·추천과 사람의 최종 지급 결정을 나누는 질문을 던진다. 아래 계획은 강의에서 제안할 실행안이다.

90일 실행 제안

기간	할 일	책임과 산출물	다음 단계로 갈 조건
1–30일	보고서 검토·승인 지침 검색·상담 후속 메모 중 한 업무 선정; 현재 작업 측정	현업 책임자·지식 소유자·검토자·IT·보안·준법 지정; 자료 범위와 검증 세트	업무·자료·도구·승인·중단 기준을 설명할 수 있음
31–60일	자료 사용이 승인된 제한 환경에서 현재 방식과 AI 지원 비교	정상·예외·구지침·권한·공격·장애 사례; 가능하면 생성자와 판정자 분리	품질·보안 기준을 유지하고 총시간·재작업을 함께 측정
61–90일	유지·확대·재설계·중단 판단	과제별 비용·오류·사고·현업 유지 공수·수동 복구 결과	현업·보안·준법이 같은 증거로 확대 범위에 합의

검증은 시험 전에 정의한다

정답과 예외, 주요 지표, 표본, 포함·제외 기준, 장애와 실패 처리 방식을 시험 전에 정한다. 비슷한 난이도의 업무를 비교하고, 같은 자료를 반복해서 생기는 숙련 효과를 줄인다. 결과를 만든 사람과 판정자를 가능한 한 분리한다. 효과는 건당 총시간 차이, 중요한 오류의 사건 수와 분모, 재작업, 현업 지식 갱신·운영 공수로 보고한다. 표본이 허용하면 짝지은 비교와 신뢰구간을 제시한다. 실험에서 위반 0건이 나와도 무위험이나 정확도 100%로 표현하지 않는다. **권한 밖 자료 노출, 고객정보 유출, 승인 없는 지급·발송·등록, 통제되지 않은 중복 실행**은 즉시 중단과 원인 확인의 조건으로 제안한다. 품질 저하, 미승인 지침 사용, 검토·유지 공수의 증가가 반복되면 확대를 보류하고 과제를 재설계한다. 이 보고서에서는 고객사 시험, ROI 계산, 인력 효과 추정, 통제 감사를 실행하지 않았다. 실행 시 회사의 기준과 업무 위험에 맞춰 검증·중단 조건을 확정해야 한다.

임원이 회의에서 결정할 다섯 가지

1. 어느 업무를 첫 과제로 선택할 것인가?
2. 판단 기준과 지식의 승인·갱신을 누가 책임질 것인가?
3. AI가 사용할 계정·자료·도구와 실행 범위는 어디까지인가?
4. 어떤 결과를 사람이 검토하고 어떤 행동을 승인할 것인가?
5. 30일 뒤 어떤 증거로 유지·확대·중단을 판단할 것인가?

9. 보고서의 범위

이 문서는 2026-10-06 기준의 기존 공개 조사와 2026-10-07 재조회 결과를 읽기 위한 보고서로 재구성했다. 공식·1차 출처 26건과 언론 보도 1건의 기능·계획·측정 범위를 사용했으며, 고객사의 비공개 운영 현황과 공개 사실을 합산하지 않았다. 기업 발표는 회사가 해당 기능·성과를 설명했다는 근거다. 운영 효과의 독립 검증이나 고객사의 준수 여부를 뜻하지 않는다. 법률·제품·승인 조건은 실제 적용 전에 해당 업무와 계정·계약에서 확인해야 한다.

10. 직접 출처

아래 각주는 기존 조사에서 대조한 직접 출처다. 회사 발표, 감독 자료, 권고, 언론 보도의 성격과 한계를 함께 남겼다.

S01 KB손해보험, AI 에이전트 사원번호 관리 (https://insight.kbinsure.co.kr/20260914-aiagent-employee/), 2026-09-14. 회사 발표의 적용 범위와 향후 평가 계획을 구분했다.

S02 KB손해보험, 언더커버 에이전터스 (https://insight.kbinsure.co.kr/260929-undercover-agenters-ai-hackathon/), 2026-09-29. 현업 개발·임원 스폰서 참여 근거이며 운영 효과 연구가 아니다.

S03 LG, LG CNS의 신한은행 금융지식 Q&A 구축 (https://lg.co.kr/media/release/28963), 2025-05-12. 공급자가 설명한 지식·RAG·보안 설계이며 독립 방어율 측정이 아니다.

S04 신한금융, 여신심사지원 Agent (https://www.shinhangroup.com/kr/archive/press/detail/679), 2026-03-27. 기업 분석과 의견서 지원 범위.

S05 카카오뱅크, 안전한 이체 활용법 (https://blog.kakaobank.com/posts/kakaobank-life-transfer-security), 2026-08-06. 공식 사용 안내의 확인·경고 동선을 참고했다.

S06 Morgan Stanley, Debrief launch (https://www.morganstanley.com/press-releases/ai-at-morgan-stanley-debrief-launch), 2024-06-26. 고객 동의·담당자 수정/발송·팀 단위 Assistant 채택과 개별 일화를 구분.

S07 DBS, 2025 CEO reflections (https://www.dbs.com/annualreports/2025/ceo-reflections.html). 2025 포트폴리오 경제가치와 조직 전환 설명; 외부 인과 검증 아님.

S08 AXA, AI workplace strategy (https://www.axa.com/press/press-releases/axa-accelerates-its-ai-workplace-strategy), 2026-07-20. Copilot 점진 배포·교육·인간 감독을 설명.

S09 Allianz, EKA (https://www.allianz.com/en/mediacenter/news/articles/250207-allianz-ai-enterprise-knowledge-assistant-enhances-productivity.html), 2025-02-07. 지식 검색·원문 연결과 제한된 회사 정확도 주장.

S10 JPMorganChase, 2024 COO shareholder letter (https://www.jpmorganchase.com/ir/annual-report/2024/ar-ceo-letter-jennifer-piepszak), 2025-04-07. 배포와 통제 플랫폼의 설명.

S11 Mastercard, Decision Intelligence Pro (https://newsroom.mastercard.com/news/press/2024/february/mastercard-supercharges-consumer-protection-with-gen-ai/), 2024-02-01. 초기 모델링 지표와 출시 예고.

S12 금융위, 모델 변경 절차 (https://www.fsc.go.kr/po010101/86712?curPage=2&srchCtgry=1), 2026-04-15. 기존 혁신금융서비스의 보안 영향별 절차.

S13 금융위, AX·AI 가이드라인 발표 (https://www.fsc.go.kr/po010101/87142), 2026-06-18. 자율규제 7원칙과 별도 법률 의무를 구분.

S14 금융위, 제2차 보안 목적 테스트 (https://www.fsc.go.kr/po010105/87646), 2026-09-03. 보안 목적·대체 통제·한시 조건·후속 예정일.

S15 금융보안원, 금융분야 인공지능 보안 안내서 (https://www.fsec.or.kr/bbs/detail?bbsNo=11977&menuNo=222), 2026-06-19 게시. 첨부 72쪽 PDF의 pp.6–7, 10–12, 28, 39–40, 56–57, 64–66 관련 범위를 대조했다. 원문 p.7도 렌더 확인. 권고와 예시는 회사별 준수 판정·법적 의무로 일반화하지 않는다.

S16 금융보안원, 프런티어 AI 보안 위협 금융분야 대응 요령 v1.0 (https://www.fsec.or.kr/bbs/detail?bbsNo=11989&menuNo=222), 2026-07-01. PDF 18쪽 중 대응 본문의 PDF pp.7–18(인쇄 pp.4–15)를 대조하고 위험별 자동화 표(PDF p.15/인쇄 p.12)를 렌더 확인했다. 비규제 안내이며 시간·배점·자동화 표는 적용 예시다.

S17 Microsoft, Copilot data, privacy, security (https://learn.microsoft.com/en-us/microsoft-365/copilot/microsoft-365-copilot-privacy), 2026-08-18 갱신. 제품 문서의 권한·학습·보관·확장 처리 조건.

S18 OWASP, LLM01:2025 (https://genai.owasp.org/llmrisk/llm01-prompt-injection/). 프롬프트 인젝션·RAG 한계·완화책.

S19 OWASP, LLM06:2025 (https://genai.owasp.org/llmrisk/llm062025-excessive-agency/). 과도한 기능·권한·자율성과 승인 통제.

S20 NIST, AI 600-1 출판 개요 (https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence), 2024-07-26. 자율적 위험관리 프로필의 성격.

S21 FinCEN, Deepfake fraud alert (https://www.fincen.gov/news/news-releases/fincen-issues-alert-fraud-schemes-involving-deepfake-media-targeting-financial), 2024-11-13. 금융기관 대상 사기 정황과 경보.

S22 EIOPA, AI governance Opinion 설명 (https://www.eiopa.europa.eu/eiopa-publishes-opinion-ai-governance-and-risk-management-2025-08-06_en), 2025-08-06. EU 보험 감독의 위험 비례 접근과 Opinion의 범위.

S23 DBS, 2024 CIO statement (https://www.dbs.com/annualreports/2024/cio-statement.html), Annual Report 2024. 상담 처리시간 20%는 기대값이며, 변경 위험평가·운영 재설계 수치는 각각 별도 업무 범위다. 2026-10-07 원문 재조회.

S24 NAIC, Members Approve Model Bulletin on Use of AI by Insurers (https://content.naic.org/article/naic-members-approve-model-bulletin-use-ai-insurers), 2023-12-04; 2026-10-07 조회. 공식 채택 설명을 확인했으며 Model Bulletin 전문·모든 주의 현행 적용을 검증하지 않았다.

S25 MAS, Project MindForge (https://www.mas.gov.sg/schemes-and-initiatives/project-mindforge), 게시 2025-11-13, 페이지 갱신 2026-03-20; 2026-10-07 조회. Phase 2·Toolkit 구성 개요이며 연결된 문서 3종의 PDF 전체 대조가 아니다.

S26 한국금융신문, AI 사이버 위협에 망분리 완화 제동…2차 선정 일정 연기 (https://www.fntimes.com/html/view.php?ud=2026100616381464062a735e27af_18), 2026-10-06 16:50; 2026-10-07 04:06 이후 직접 본문 조회. 보도·관계자 인용과 공식 새 일정/발급 상태를 구분한다.

S27 Allianz, When the storm clears, so should the claim queue (https://www.allianz.com/en/mediacenter/news/articles/251103-when-the-storm-clears-so-should-the-claim-queue.html), 2025-11-03. 공개 웹 검색이 반환한 공식 본문을 대조했으며 직접 HTTP는 403이었다. 제한 청구 범위·사람 지급 결정과 회사 성과 주장을 구분한다.

이 보고서가 주장하지 않는 것

- 회사 발표 수치로 KB손해보험의 ROI·인력 대체·손해율 개선·보안 방어율을 확정하는 주장은 기각했다.
- 해외의 제품·규제·사례를 고객사의 현재 사용 승인으로 바꾸는 해석은 기각했다.
- AI 제품명과 'Agent'라는 명칭만으로 자동 지급·완전 자율 실행을 추정하지 않았다.
- 회사별 효과의 독립 검증, 실제 계정·계약·통제·승인 상태, 시험 결과는 확인되지 않은 상태로 남긴다.