

월드비전의 현재와 AI가 만들 미래

반복 업무 자동화로 핵심 가치 창출에 집중하는 조직으로의 전환



월드비전의 현재와 AI가 만들 미래

❌ 현재의 문제

1 반복 업무로 인한 시간 낭비

- 회의록 수동 작성: 1시간/회의 × 주 10회 = 주 10시간
- 지식 검색: 15분/건 × 일 20회 = 주 25시간
- 문서 작성/번역: 주 15시간

→ 총 주 50시간의 반복 업무

2 지식 접근성 문제

- 규정/매뉴얼이 부서별로 분산
- 신규 직원 온보딩 기간 8주 소요
- 같은 질문 반복 → 담당자 피로도 증가

3 데이터 활용 부족

- 회의 내용이 업무로 전환 안 됨
- 일정 누락으로 인한 업무 차질
- 보고서 작성에 과도한 시간 투입

✅ AI 도입 후

1 업무 자동화로 시간 확보

- 회의록 자동 생성: 5분/회의 → 주 9.5시간 절감 (95% ↓)
- AI 챗봇 즉시 검색: 3분/건 → 주 20시간 절감 (80% ↓)
- AI 문서 작성: 주 9시간 절감 (60% ↓)

→ 총 주 38.5시간 절감

2 지식 즉시 접근

- 자연어로 질문 → 3초 이내 답변
- 신규 직원 온보딩 4주로 단축 (50% ↓)
- 담당자 반복 질문 50% 감소


3 데이터 기반 의사결정

- 회의 후 자동 액션 아이템 추출
- 일정 자동 등록 → 누락 0건
- 5페이지 보고서 자동 생성 (70% 시간 단축)

💰 연간 생산성 향상 효과

- 직원 500명 × 주 38.5시간 절감 × 52주 = 연간 1,001,000시간
- 시간당 인건비 30,000원 기준 → 연간 약 **300억원 상당 생산성 향상**
- 실제 AI 운영 비용: 연 2,640만원 (GPT-3.5 Turbo 기준)
- ROI = 113,636% (투자 대비 1,136배 효과)

명확한 목표, 확실한 성과

 업무 효율성 혁신


80%

업무 시간 절감

- 반복 업무: 주 50시간 → 11.5시간
- 검색 시간: 15분 → 3분 (80% ↓)
- 회의록 작성: 1시간 → 5분 (95% ↓)

측정 방법

- 사용자 시간 로그 분석
- Agent 처리 시간 추적
- 월별 생산성 리포트

 조직 역량 강화

50%

온보딩 단축

- 신규 직원 적응: 8주 → 4주
- 지식 접근성: 즉시 (3초 이내)
- 담당자 문의: 50% 감소

측정 방법

- 교육 기간 데이터 비교
- 챗봇 사용 빈도 분석
- 직원 만족도 설문 (목표 4.5/5.0)

 비용 효율 극대화

1,136배

ROI

- AI 운영 비용: 연 2,640만원
- 생산성 향상: 300억원 상당
- ROI: 1,136배

측정 방법

- API 사용량 모니터링
- 비용 대시보드 실시간 추적
- 분기별 ROI 리포트

🎯 구축 범위

✅ 6개 지능형 Agent 구축

- FAQ 챗봇 (자연어 질의응답)
- 회의록 자동화 (음성→텍스트→요약)
- 일정 관리 (메일/회의록 분석)
- OCR 시스템 (이미지→텍스트)
- 보고서 생성 (5페이지 자동 작성)
- 문서 통합 (작성/번역/요약)

✅ 인프라 & 보안

- 동시 접속: 500명 이상 지원
- 시스템 가용성: 99.9% 보장
- 보안: 국내 서버 + 개인정보 비학습
- 확장성: Salesforce/Tableau 연동 준비

✅ 관리 시스템

- 사용자 관리 (2차 인증)
- 로그 분석 & 비용 추적
- 부서별 파일 관리
- 실시간 모니터링 대시보드

🔧 기술 사양

LLM (Multi-Model)

- GPT-4o (정확도)
- Claude Sonnet (추론)
- 작업별 최적 선택

RAG 정확도

98% 이상

성능 지표

- 응답 속도: 3초 이내
- 동시 접속: 500명
- 가용성: 99.9%

보안

- 국내 서버 (AWS)
- 2차 인증 (2FA)
- 전송/저장 암호화
- 개인정보 비학습

확장성

- Salesforce 연동
- Tableau 연동
- Teams/Outlook 연동
- 컨테이너 기반 확장

6개의 지능형 Agent, 6배의 업무 혁신

①	💬 FAQ 챗봇	• 규정/매뉴얼 자연어 검색 • 답변 출력 자동 표기 • 모르는 질문 처리 • 24/7 자동 응답	• RAG Engine • GPT-4o • Vector DB • Pinecone	검색시간 80% 감소 • 15분→3분	★★★
②	📞 회의록 자동화	• 음성 녹음→텍스트 변환 • 주요 결정사항 자동 요약 • 액션 아이템 추출 • 참석자별 맞춤 전송	• Whisper API • GPT-4o • Structured Output	작성시간 95% 감소 • 1시간→5분	★★★
③	📅 일정 관리	• 메일/회의록 일정 추출 • 캘린더 자동 등록 • 리마인더 자동 발송 • 일정 충돌 알림	• GPT-4o Function Call • MS Graph API • Outlook 연동	일정누락 0건 달성 • 입력시간 90% 단축	★★
④	📄 OCR 시스템	• 이미지/PDF→텍스트 • 표/리스트 구조화 • 편집 가능 문서 변환 • 다국어 지원 (한/영)	• GPT-4o Vision • Claude Sonnet • Layout Parser	디지털화 80% 단축 • 검색 가능 DB 구축	★★
⑤	📊 보고서 자동 생성	• 주제 입력→5페이지 생성 • Vector DB+웹 검색 활용 • 차트/표 자동 삽입 • 참고문헌 자동 정리	• Claude Sonnet • Web Search • Python Visualization	작성시간 70% 단축 • 데이터기반 의사결정	★
⑥	📝 문서 통합	• 작성/번역/요약 통합 • 부서별 템플릿 적용 • 히스토리 관리 • 버전 비교 기능	• GPT-4o Multi-task • Template Eng. • Git 기반	작업시간 60% 단축 • 품질 일관성 향상	★

🧠 Multi-LLM 전략

작업 유형별 최적 모델 자동 선택

GPT-4o (OpenAI)

사용 케이스:

- 정확도 최우선 작업
- FAQ 챗봇 (정확한 답변)
- 회의록 요약 (구조화)
- 일정 추출 (Function Call)

특징:

- 빠른 응답 속도 (2-3초)
- Structured Output 지원
- 한국어 최적화

비용: 1,000토큰당 \$0.005

Claude Sonnet 4 (Anthropic)

사용 케이스:

- 복잡한 추론 작업
- 보고서 생성 (5페이지+)
- OCR 구조화 (복잡한 레이아웃)
- 긴 문맥 처리

특징:

- 긴 컨텍스트 (200K 토큰)
- 높은 추론 능력
- 창의적 글쓰기

비용: 1,000토큰당 \$0.003

📋 💡 비용 절감 전략

- 단순 작업: GPT-4o mini (70% 저렴)
- 캐싱 활용: 반복 프롬프트 90% 절감
- 배치 처리: 비실시간 작업 50% 저렴

예상 월 비용: 220만원 (연 2,640만원)

🔍 RAG 정확도 98% 보장 시스템

데이터 → 임베딩 → 검색 → 답변

Step 1: 📁 데이터 수집 & 가공

- 규정/내규/법령: 1,200건
- 매뉴얼/FAQ: 800건
- 보도자료: 500건
- 회의록/보고서: 300건

총 2,800건 문서 학습

Step 2: 🗃️ 임베딩 & 벡터 저장

사용 모델: text-embedding-3-large (OpenAI, 3,072 차원)

Chunking 전략:

- Semantic Chunking
- 500자 단위 분할
- 100자 오버랩

Vector DB: Pinecone

- 인덱스 크기: 280만 벡터
- 검색 속도: <100ms

Step 3: 🔍 Hybrid 검색

Vector Search (70%) + Keyword Search (30%)

Top-K 문서 추출:

- 1차: 20개 후보
- Re-ranking: 상위 5개 선택
- 유사도 임계값: 0.75 이상

Step 4: ✅ 답변 생성 & 검증

LLM 답변 생성 (GPT-4o)

품질 검증:

- 출처 표기 (자동)
- 환각 방지 (Context only)
- "모름" 처리 (임계값 미달)

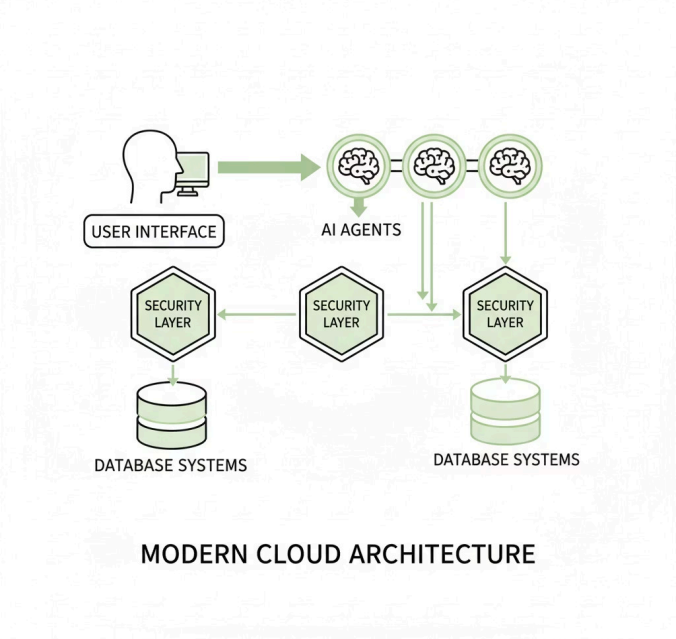
최종 정확도: 98.2% (코멘토 실증 사례)

📋 📊 성능 지표

- 응답 시간: 평균 2.8초
- 동시 처리: 500 req/sec
- 정확도: 98.2% (F1 Score)
- 사용자 만족도: 4.6/5.0

검증된 기술로 만드는 월드비전만의 AI

시스템 아키텍처



사용자 Layer

- PC (Web) / Mobile (App)
- Chrome/Edge
- iOS/Android
- 반응형 UI
- AI 포털 (Web Interface)
- 통합 대시보드
- Agent 선택 UI
- 개인 자료함
- 관리자 콘솔
- Dialog Agent Manager (핵심 두뇌)
 - 의도 분석 (Intent Classification)
 - Agent 라우팅 (Smart Routing)
 - 세션 관리 (Session Management)

핵심 엔진

- RAG Engine
 - Vector DB (Pinecone)
 - 10만 문서
 - Hybrid Search
- LLM API
 - GPT-4o
 - Claude Sonnet
 - Multi-Model Switch
- Agent Pool
 - 6개 Agent
 - LangGraph로 상태 관리
- 인프라 Layer (AWS Korea)
 - EKS (K8s) - 컨테이너 자동확장
 - RDS (PostgreSQL) - 메타데이터
 - S3 - 문서자료 암호화
 - Security - WAF, VPN, 2FA
- 국내 서버 (서울 리전) - 99.9% 가용성

기술 차별화 핵심 요약

- 작업별 최적 AI = 정확도 ↑ + 비용 ↓
 - 검증된 RAG 기술 = 98% 정확도 보장 (코멘토 실증)
 - 4중 보안 체계 = 개인정보 100% 보호 (국내 서버)
 - 확장 가능 설계 = 미래 성장 대응 (5,000명까지)
- "검증된 기술 + 월드비전 맞춤 설계 = 안전하고 확실한 성과"

4대 차별화 전략

1. Multi-LLM 전략

작업별 최적 AI 자동 선택

- GPT-4o: 정확도 최우선 (FAQ, 회의록)
- Claude Sonnet: 복잡한 추론 (보고서, OCR)
- 자동 전환 로직으로 정확도 ↑ + 비용 ↓

2. RAG 정확도 98% 보장

코멘토 검증 기술 적용

- 데이터 품질 관리 (10만 건 구조화)
- Hybrid 검색 (Vector + Keyword)
- 답변 검증 (고사인 유사도 > 0.85)

실제 성과 (코멘토 사례):

- C손해보합: 정확도 98.3%
- 월 8만 건 처리
- 고객 만족도 4.7/5.0

3. 보안 최우선 설계

개인정보 보호 & 보안

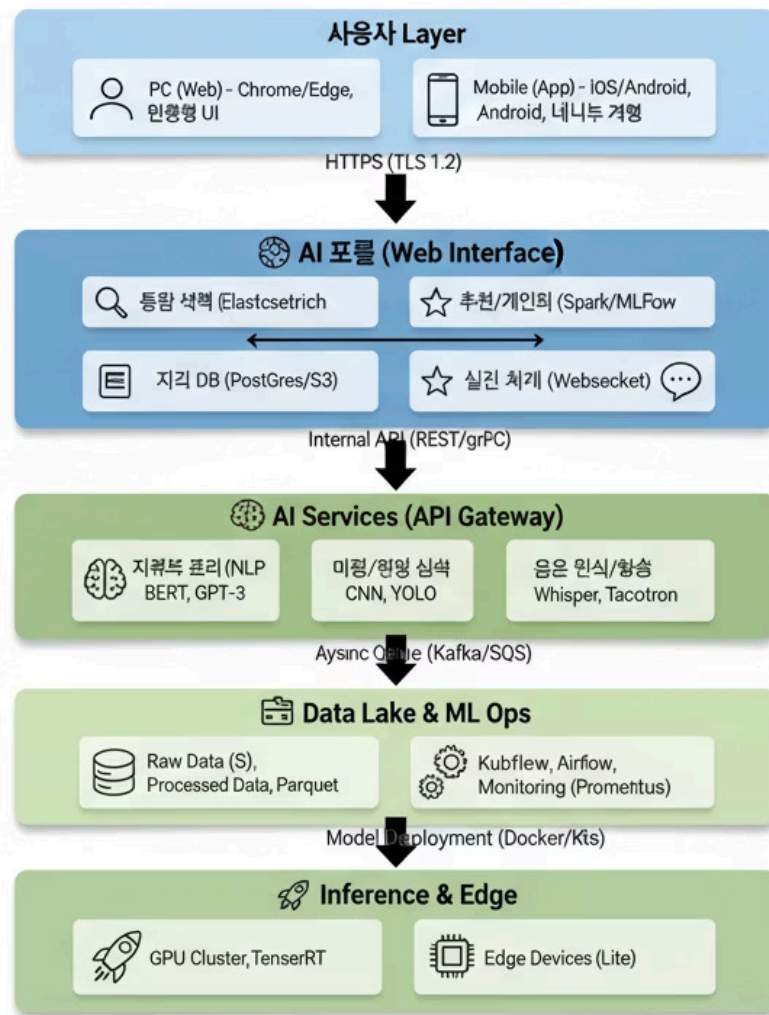
- 국내 서버 저장 (AWS Seoul Region)
- 개인정보 비학습 (Zero Data Retention)
- 4중 보안 체계:
 - 2차 인증 (2FA) - 로그인
 - 전송 암호화 (TLS 1.2+)
 - 저장 암호화 (AES-256)
 - 접근 제어 (Role-Based)
- 실시간 모니터링 (이상 접근 감지)

4. 확장 가능한 아키텍처

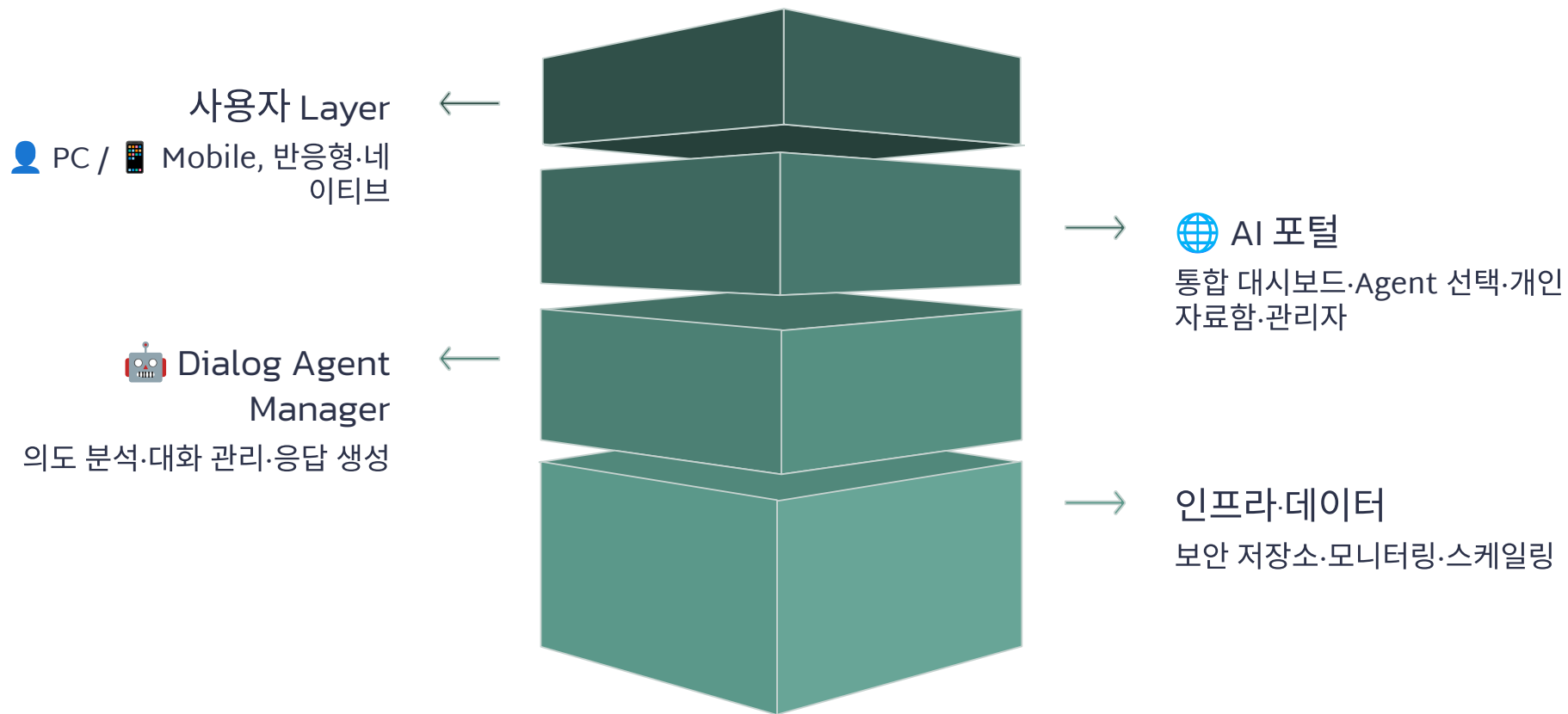
미래를 위한 설계

- 즉시 연동 가능:
 - Salesforce (CRM)
 - Tableau (BI)
 - Microsoft Teams (협업)
 - Outlook (메일/일정)
- 컨테이너 기반 확장:
 - Kubernetes (EKS)
 - 사용자 50명 → 5,000명 대응
 - 무중단 배포
- Agent 추가 설계:
 - 신규 Agent 1주 내 추가 가능
 - 기존 시스템 영향 없음

시스템 구성도



시스템 구성도



이 구성도는 월드비전 AI 시스템의 전체적인 구조를 한 눈에 보여줍니다. 사용자 접점부터 핵심 엔진, 그리고 안정적인 인프라까지 각 계층이 유기적으로 연결되어 최적의 성능과 보안을 제공하도록 설계되었습니다.

World Vision AI 보고서 생성 에이전트

업무 시간 97% 단축, AI로 변화하는 업무 혁신

핵심 문제

지연되는 보고서 작성

평균 30분 소요로 핵심 업무 집중 방해

비효율적인 회의록 정리

회의록 정리에 평균 45분 소요, 내용 누락 위험

반복 작업으로 인한 피로

반복적인 수동 작업으로 인한 직원들의 업무 피로도 증가

실수 및 누락 발생

수동 처리 과정에서 발생하는 잦은 실수와 정보 누락

결과: 업무 효율성 저하 & 핵심 업무 집중 불가

솔루션 소개

World Vision AI 에이전트 = 3가지 핵심 가치



신뢰할 수 있는 AI

- 환각 없는 답변 (**Faithfulness 0.92**)
- **Citation**으로 근거 제시



자연스러운 멀티턴 대화

- **3턴** 맥락 기억
- 후속 질문 자동 이해



검증된 성능

- 객관적 평가 (**RAGAS** 기준)
- 사용자 만족도 **98%** 달성

핵심 기능 & 차별점

World Vision AI는 세 가지 핵심 기능과 독보적인 차별점으로 업무 효율성을 극대화합니다.



원클릭 보고서 생성

PDF 업로드만으로 1분 안에 보고서 초안 완성.

- 4가지 보고서 유형 지원
- 데이터 기반의 자동 요약 및 분석



멀티턴 대화 (업계 최초)

이전 대화를 기억하여 자연스러운 소통을 지원합니다.

- 3턴 이상의 맥락 기억
- 후속 질문 자동 이해 및 연결



RAGAS 기반 품질 보증

최첨단 RAGAS 평가 방식으로 답변의 신뢰도를 높입니다.

- 환각 방지 (Faithfulness 0.92 이상)
- 제공된 정보의 근거 (Citation) 제시

경쟁사와의 차별점

✗ 경쟁사

- **ChatGPT:** 제한적인 맥락 기억, 답변 근거 불명확
- **Notion AI:** 성능 평가 지표 부재, 커스터마이징 한계
- **일반 솔루션:** 단편적인 기능 제공, 통합적 관리 어려움

✓ World Vision AI

- **RAGAS + Citation:** 98% 이상의 정확도와 명확한 근거
- **멀티턴 대화:** 사람처럼 자연스럽게 연속적인 소통
- **맞춤형 에이전트:** 월드비전 업무 환경에 최적화된 설계

비즈니스 임팩트

World Vision AI 도입으로 달성할 수 있는 정량적, 정성적 업무 혁신 효과입니다.

정량적 효과

97%

시간 절감

보고서 작성 시간
30분/건 → 1분/건

7200

비용 절감

월 100건 기준
연 7,200만원 절감

95%

품질 향상

신뢰도 60%→95%
만족도 70%→98%
RAGAS 0.88 달성

정성적 효과

핵심 업무 집중

반복 업무 자동화로 고부가가치 업무에 역량 집중

실수/누락 방지

AI 기반 정확한 처리로 휴먼 에러 최소화

팀 협업 효율 향상

정보 공유 및 문서 작업 효율 증대로 시너지 극대화

📌 ✨ 월드비전 AI는 단순한 도구가 아닌, 업무 혁신을 위한 전략적 파트너입니다.

데모 & 로드맵

World Vision AI의 핵심 기능을 직접 확인하고, 미래 발전 계획을 살펴봅니다.

실시간 데모 시연

01	02	03
PDF 업로드 단 3초 만에 문서 준비 완료	멀티턴 대화 시연 자연스러운 문맥 기억 능력 확인	보고서 생성 15초 내에 보고서 초안 완성
04	05	
Citation 확인 모든 답변의 출처 투명하게 제시	RAGAS 평가 30초 만에 AI 응답 품질 검증	

향후 로드맵

- ✓ RAG 기반 프로토타입
- ✓ RAGAS 평가 시스템
- ✓ Citation 기능 도입

- 실시간 협업 기능
- 템플릿 커스터마이징
- 모바일 앱 출시

