고객 금융 이해력 분석에 따른
생성형 AI 기반 개인화 번역 서비스

FINTERPRET

작성자
조효원, 이채훈

제출일
2026.01.09

주최
한화생명

INDEX
목차

01 시장 현황과 구조적 문제 — 3p-7p

02 문제 정의와 기획 목표 — 8p

03 서비스 컨셉 및 핵심 아이디어 — 9p-10p

04 사용자 시나리오 분석 — 11p-13p

05 차별성 및 한계 분석 — 14p-16p

06 기술적 구현 — 17p-21p

07 기대 효과 및 사회적 가치 — 22p

08 비즈니스 모델 및 확장 로드맵 — 23p-25p

시장 현황: 금융 문맹의 구조적 문제

모두를 문맹으로 만드는 금융 계약서

국내 금융 이해력 실태

금융 이해력이 낮으니 약관 읽기도, 상품 선택도 어렵다

- 국내 성인의 평균 금융 이해력은 중간 수준에 머물러 있으며, 복합 금융상품(보험·파생상품)으로 갈수록 이해도가 급격히 하락
- 특히 장기 재무계획, 위험–수익 구조, 조건부 지급에 대한 이해가 취약
- 다수의 소비자가 금융상품 선택 시 약관 이해보다 판매 설명·주변인 추천에 의존
- 이는 금융 소비자의 문제가 아니라,이해를 전제로 설계되지 않은 금융 설명 구조의 문제

출처 | 한국은행·금융감독원, 2024 전국민 금융이해력 조사 결과 보도자료 (2025.04.29)

보험약관 이해도 평가

보험사도 보험약관 이해도 향상에 어려움을 겪는다

- 보험약관 이해도 평가는 2011년 도입 이후 지속되어 왔으나, 대다수 보험사의 약관은 여전히 '보통~미흡' 수준
- 약관 분량은 길고, 법률·의학·금융 용어가 혼재되어 있어 이해도 향상이 쉽지 않은 구조
- 특히 면책 조항, 예외 규정, 지급 제한 조건에서 소비자 이해도가 급격히 하락
- 제도적 평가는 존재하지만, 실제 소비자의 이해를 높이는 구조적 개선은 제한적

출처 | CEO스코어데일리, 보험약관 이해도 평가 결과 분석 기사

왜 금융 문맹 보완이 중요한가?

금융 문맹은 개인의 손해를 넘어, 시장 신뢰를 훼손한다

- OECD 및 유럽은행감독청(EBA)은 금융 계약서 이해 부족의 문제가 다음과 같은 문제로 이어짐을 지적

  문제 a. 불완전 판매

  문제 b. 민원·분쟁 증가

  문제 c. 금융시장에 대한 불신 확대로 연결
- 금융소비자보호법 이후에도 여전히 금융분쟁 조정 민원은 증가하고 있음. 수십만 건의 금융분쟁 민원은 사실상 방치되고 있어 금융 소비자 보호가 제대로 되고 있지 않음

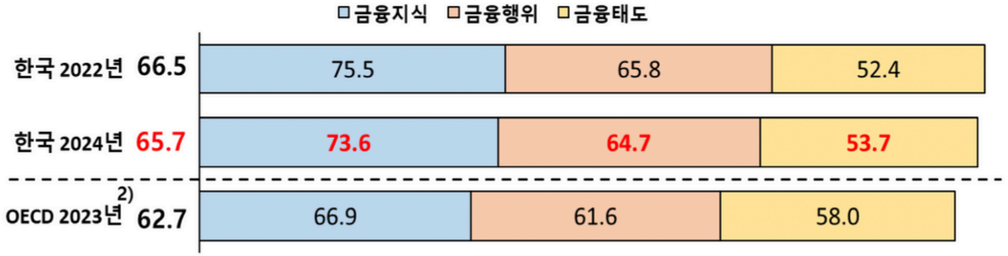
출처 | OECD & IOSCO, 금융 문해력·행동 경제학 보고서
출처 | EBA FINANCIAL EDUCATION REPORT 2019/2020
출처 | 조선비즈, 금융분쟁조정 현황 분석 기사 (2025.08.26)

시장 현황: 금융 문맹의 구조적 문제

금융소비자보호법 시행을 통한 설명 의무 제도화와 여전한 이해 검증의 한계

낮은 점수로 정체된 금융이해력, 가장 취약한 세대는 20대
금융이해력의 공백은 이미 구조적인 문제로 고착화되고 있다

우리나라 성인 금융이해력 점수[1)]



주: 1) 금융이해력 점수 = 금융지식×35% + 금융행위×45% + 금융태도×20%
2) OECD 국가중 금융이해력 조사에 참여한 20개국 점수의 단순평균

응답자 특성별 금융이해력

	연령						소득계층[1)]			학력		
	20대[2)]	30대	40대	50대	60대	70대	저소득	중소득	고소득	고졸미만	고졸	대졸[3)]이상
2022년 (A)	65.8	69.0	68.9	67.0	64.4	61.1	63.2	68.0	68.7	59.3	65.4	68.7
2024년 (B)	62.6	67.7	68.4	67.9	64.7	59.3	59.7	66.8	68.8	59.0	64.5	68.0
차이 (B-A)[4)]	-3.2	-1.3	-0.5	0.9	0.3	-1.8	-3.5	-1.2	0.1	-0.3	-0.9	-0.7

주 : 1) 연소득 7천만원 이상 고소득층, 3천만원~7천만원 중소득층, 3천만원 미만 저소득층으로 구분
2) 만 18~29세 3) 전문대학 포함 4) 푸른색 음영은 22년 대비 하락, 주황색 음영은 22년 대비 상승한 계층

출처 | 한국은행·금융감독원, 2024 전국민 금융이해력 조사 결과 보도자료 (2025.04.29)

금융분쟁 민원은 해마다 증가하고 있으나 실제로 분쟁조정위원회에 회부된 사례는 적으며, 이는 금융 소비자 보호 체계에 구조적 허점이 존재함을 시사하며 사회적 비용을 초래한다

연도별 금융분쟁 민원 및 분쟁조정위원회 회부 건수

출처 | 조선비즈, 금융분쟁조정 현황 분석 기사 (2025.08.26)

설명은 되었지만, 이해는 되지 않았다

기존 금융 약관 설명 구조 분석을 통한 이해 난이도 발생 요인 도출

기존 금융 약관·설명서의 구조적 복잡성

과도한 분량

- 보험·투자 상품 약관은 평균 수만 자에 달함
- 소비자가 계약 전 전부 읽고 이해하기 어려운 구조
- ‘중요 내용’과 ‘법적 면책 문구’가 동일한 밀도로 나열됨

법률·기술 용어 중심 서술

- 법적 분쟁을 대비한 표현 위주
- 일상 언어가 아닌 법률·의학·금융 전문 용어 혼합
- 평균 소비자 독해 수준을 초과하는 문장 구조

조건부 구조의 복합성

- “보장되나, 특정 조건에서는 지급 제외”
- 면책·예외·공제·감액 조건이 다층적으로 얽힘
- 소비자는 ‘보장’과 ‘지급 가능성’을 동일시하는 경향

설명 방식의 획일성

- 모든 소비자에게 동일한 설명 자료 제공
- 금융 이해도·경험 차이를 고려하지 않음
- 실제 금융 리스크는 여러 변수 (연령대, 자산 규모, 투자 경험, 재무 여력)에 따라 의미가 달라짐

AI 금융 서비스도 '이해'를 시키진 못한다

기존 AI 기반 금융 약관 설명 시스템 유형 분석 및 한계점 정리

현존 AI 금융 서비스의 한계

요약 중심 AI

요약은 되지만, 이해는 보장되지 않는다.

- 약관을 줄이거나 핵심 조항을 발췌하는 데 초점
- 설명 난이도는 그대로 유지
- 사례: KB손해보험 보험상품 AI 에이전트, 신한라이프 AI 설계 에이전트

Q&A 챗봇 중심 AI

질문하지 않으면, 위험도 보이지 않는다.

- 고객은 무엇을 모르는지 인지해야 질문할 수 있으나, 실제 분쟁을 야기하는 면책·예외·조건부 지급 조항은 '모른다는 사실조차 인식되지 않는 영역'에 놓여 있음
- 질문되지 않은 면책·예외·조건부 지급 리스크는 드러나지 않음
- 사례: KB손해보험·신한라이프 자연어 기반 AI 상담

이해 검증 부재

설명은 자동화됐지만, 이해는 검증되지 않는다.

- 설명 여부는 기록되나, 이해 여부는 확인되지 않음
- 금융소비자보호법의 '이해 중심 보호' 취지와 괴리 발생
- 사례: 교보생명 보장분석 AI, 한화생명 가입설계 AI

이해의 공백이 낳은 비용: 불완전 판매와 분쟁

금융소비자보호법 제정 이후 분쟁 사례 분석을 통한 구조적 공방 문제 확인

이해의 공백을 해소하기 위해 대상의 금융 이해 수준(User Level)에 따라 금융 약관 또는 상품 설명서를 쉬운 언어로 '번역'하여 제공하는 시스템의 필요성이 제기된다.

이해의 공백은 '소비자' 뿐만 아니라 상품을 직접 설명·권유하는 '보험설계사(보험모집인)'에게도 존재한다.

문제점 및 개선사항 요약표

구분	사례 개요	분쟁 발생 원인	개선사항
불완전 판매 사례 (보험설계사 교육/관리 문제)	보험설계사가 보험사로부터 받은 교육 및 안내자료를 바탕으로 상품을 설명하였으나, 실제 약관상 지급 요건과 차이가 발생하여 보험금 지급 단계에서 분쟁으로 이어짐	보험사 교육이 요약·판매 중심으로 이루어져 약관의 구조적 의미와 지급 제한 조건에 대한 이해가 부족함. 설계사 역시 약관을 정확히 해석하기 어려운 상태에서 설명이 이루어짐	보험설계사 교육을 보조하는 약관 '번역' 시스템 필요. 약관을 설계사 이해 수준에 맞게 구조적으로 해석·전달하여, 설명 단계에서의 오인과 불완전 판매를 사전에 차단
약관 이해 실패 사례 (소비자 문제)	소비자가 약관을 읽고 자신의 상식에 따라 보험금 지급이 가능하다고 이해했으나, 법원은 약관의 문언·체계 해석을 기준으로 보험사 손을 들어주어 소비자 패소	약관은 존재했으나 법률·의학·보험 용어가 혼재되어 일반 소비자가 정확한 해석에 도달하기 어려움. 이해 여부를 검증하는 절차 부재	소비자의 금융 이해력에 맞춰 약관을 의사결정 관점에서 번역하고, 분쟁 가능성이 있는 해석을 사전에 경고하는 개인화된 설명 도구 필요

출처 | 대법원, 보험설계사의 설명의무 위반으로 보험사 손해배상 책임 인정 판결 (대법원 2022다200317·200324, 2024.12.12)

목표: 언어의 장벽을 허무는 금융 인프라

교육·설명 구조의 한계를 통한 기술 기반 보완 필요성 제기

- 금융 정보의 비대칭은 정보량이 아닌 언어와 난이도에서 발생한다.

Mission

- 금융 소외 계층과 공급자 사이의 언어적 장벽을 기술로 철거.
- 누구나 동등한 정보력을 갖는 금융 환경 구축.

Vision

1. 소비자 주권
- 모르고 서명하는 공포해소.
- 금융 이해력(Literacy) 격차 제로(0)화.

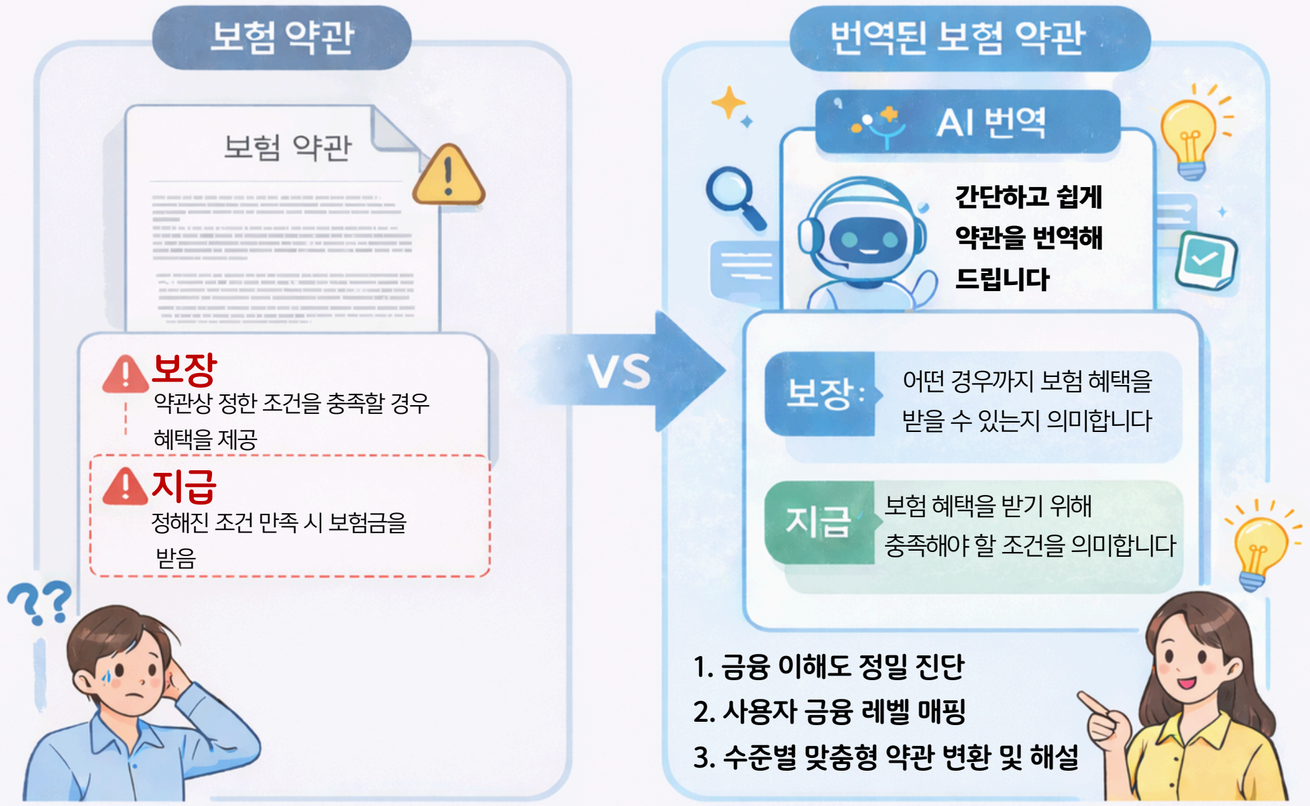
2. 기업의 건전성
- 불완전 판매율 0% 도전.
- ‘설명'이 아닌 ‘이해’ 중심 프로세스 정착.

3. 사회적 신뢰
- “설명했다 vs 못 들었다” 공방 종식.
- 데이터 기반 투명한 금융 분쟁 해결 문화 조성.

FINTERPRET: 생성형 AI 기반 개인화 번역 서비스

문서 입력 기반 개인화 번역 구조를 통한 서비스 핵심 아이디어 도출

본 서비스에서 말하는 '번역'의 정의란, ① 법적 효력은 유지 ②위험·조건·손익 구조를 사용자의 금융 문해력(User Level)에 맞춰 재구성하여 설명의 편의가 아니라 의사결정 품질을 높이기 위한 설계이다.
FINTERPRET는 기존 요약 및 질의응답 형식의 AI 챗봇 상담 시스템이 아닌, 상담 이전 단계에서 이해와 계산을 도와주는 독립 엔진으로 작동한다.



금융 문해력 진단 (Dual Track Diagnosis)

Track A (Passive):

마이데이터 연동
(투자 경험, 보유 자산 종류, 연령대 분석)
→ 별도 입력 없이 자동 추론

Track B (Active):

최초 1회, 3~5 문항의 금융 상식 퀴즈

예시) Q. 다음 중 '원금이 보장되지 않을 수 있는 상품'은 무엇일까요?
A. 정기예금
B. 적금
C. 주식
D. 보통예금

AI가 투자 경험(Track A)과 상식 퀴즈(Track B)를 가중 합산해 금융 문해력 점수를 계산하고, 합산 점수에 따라 사용자 레벨을 자동 매핑한다.

매핑된 유저 레벨에 따라 맞춤 개인화하여 번역된 금융 약관 혹은 상품 설명서를 제공한다

FINTERPRET 설계: AI Risk Sentinel

위험 요소 선제 노출 방식을 통한 금융 약관 이해 효율성 제고 방안

사용자의 금융 수준 문해력에 맞춰 의미를 재구성하는 개인화 번역 시스템

먼저, 서비스는 '진단 → 매핑 → 번역'의 3단계 구조를 통해 사용자의 금융 이해력을 정밀하게 파악한다. 금융 문해력 진단은 마이데이터를 활용한 Passive Track과 간단한 퀴즈 기반의 Active Track을 병행하여, 별도의 부담 없이도 사용자 수준을 추론할 수 있도록 설계되었다. 이렇게 도출된 사용자 수준은 Lv.1부터 Lv.5까지로 구분되며, 금융 입문자부터 전문 투자자까지 각 단계별로 다른 설명 전략이 적용된다. 이후 시스템은 해당 수준에 맞춰 금융 약관의 핵심 의미를 번역하되, 법적 효력을 훼손하지 않기 위해 보조 역할만 수행하며 '원문'과 '원문 대비표'가 반드시 함께 제시된다.

특히, 본 서비스의 핵심 기능인 AI Risk Sentinel은 약관 전체를 문맥 단위로 분석해, 소비자에게 불리하게 작용할 수 있는 '독소 조항'을 자동으로 탐지한다. 이는 단순한 키워드 검색이 아니라, 원금 손실 조건, 중도 해지 수수료, 변동 금리 및 자동 갱신 조항 등 실제 의사결정에 영향을 미치는 조건을 중심으로 작동한다. 탐지된 위험 요소는 직관적 UI로 시각화되어, 위험도에 따라 각각 초록(안전), 노랑(주의), 빨강(위험)의 배경색으로 표기된다. 따라서, 사용자가 복잡한 수치를 해석하지 않더라도 위험 수준을 즉시 인지할 수 있다.

단순 키워드 매칭이 아닌, 문맥(Context) 분석을 통해 사용자에게 불리한 조건을 식별

직관적 시각화 (Traffic Light UI): 복잡한 수치를 위험도에 따라 색을 다르게 표기

- 초록 (안전): 예금자 보호 대상, 확정 금리.
- 노랑 (주의): 실적 배당형, 3년 내 해지 시 수수료 발생
- 빨강 (위험): 원금의 최대 100% 손실 가능 (가장 눈에 띄게 강조)

리스크 번역 (Risk Translation): 약관 문구 vs AI 문구를 비교해서 보여줌

보험설계사 관점 서비스 활용 시나리오

기존 AI 상담 이전 단계에서 이해와 계산을 보완해주는 독립 엔진 활용 시나리오 (대화형 UI는 입력 방식 중 하나)

① 금융 이해력 진단 서비스

현재 고객을 만나, 금융 약관을 설명드리기 전 단계입니다.
우선, Track A(마이데이터)와 Track B(금융 상식 퀴즈)
를 활용하여 고객의 금융 이해력을 진단해주세요.

Track A 분석을 통해 고객님이 주로 접해온 금융
영역을 확인합니다.

Track B 분석을 위해 고객님에게
금융 상식 퀴즈 3개를 제안드리겠습니다.

② User Level별 번역 서비스

Dual Track Diagnosis 분석 결과,
고객님의 금융 이해력(User Level)은
Lv.3 (일반 투자자) 수준으로 추정됩니다.

해당 Level에 알맞게 첨부한 금융 약관을
번역해줘.
<첨부파일: 금융 약관 설명서.pdf>

User Level에 알맞게 해당 금융 약관을
번역해드렸습니다. PDF 파일을 다운받아주세요.
<첨부 파일: 번역 버전 금융 약관.pdf>

너의 번역 통해서 설명 잘 마쳤어.
이제 AI 상담 서비스를 작동해서,
고객의 질문에 친절히 응답해줘.
(기존 질의응답 시스템으로 연계)

AI 번역이 해결하는 지점 시나리오 분석

리스크 가시화 및 언어 전환을 통한 실질적 설명 의무의 기술적 구현 가능성

1. 원문 약관

“본 계약은 암 치료를 보장하나, 약관 제○조에 따라 일부 치료는 지급 대상에서 제외될 수 있음”

→ ‘지급 조건’은 뒤쪽 약관에 숨겨져 있어 법적으로는 정확하지만, 소비자 입장에서는 치명적인 정보 비대칭이 발생한다

2. 소비자 인식 흐름 시나리오

1. “암 치료 보장”이라는 사고 유형만 기억함
2. “일부 치료 제외”는 추상적 문구로 인식
3. ‘지급 조건’이 어디에 있고, 무엇인지 알지 못함
4. 실제 치료 후 보험금 미지급
   → “설명 못 들었다” 분쟁 발생

→ 소비자는 ‘무엇이 지급되지 않는지'를 상품 구매 당시 질문조차 하지 못한다.
(자신이 무엇을 모르는지 모르기 때문)

AI 번역이 해결하는 지점 시나리오 분석

AI 번역 적용 시나리오 분석을 통한 기존 설명 방식과의 차별성 도출

사용자 금융 레벨별 원문 약관 번역 차이

Lv.1

금융 입문자

이 보험은 '암 치료'라는 사고를 보장하지만, 암과 관련된 모든 치료가 다 보험금으로 나오는 것은 아닙니다. 어떤 치료가 지급되는지는 약관에 정해진 조건을 확인해야 합니다.

⬇

✓ 사고(보장) vs 돈(지급) 완전 분리
✓ 판단 문장을 결론부터 제시

Lv.2

금융 관심층

암 치료는 보장 범위에 포함되지만, 치료 방법이나 시점에 따라 보험금이 지급되지 않는 경우도 있습니다. 이는 약관에서 정한 지급 조건에 따라 달라집니다.

⬇

✓ '왜 안 나오는지'를 한 단계 더 설명
✓ 조건이 존재함을 인지시킴
✕ 조항 번호·전문 용어는 아직 미사용

Lv.3

일반 투자자

암 진단 및 치료는 보장 항목에 해당하나, 약관 제○조에 따라 일부 치료는 지급 대상에서 제외될 수 있습니다. 따라서 지급 여부는 해당 치료가 지급 요건을 충족하는지에 따라 결정됩니다.

⬇

✓ 보장 / 지급 / 제외 구조 명확화
✓ 조항 번호 제시 시작
✓ 표준 금융 용어 사용

Lv.4

적극 투자자

본 계약은 암 치료를 보장하지만, 약관 제○조 ○항에서 정한 지급 요건을 충족하지 않는 치료는 보장 대상이더라도 보험금 지급이 제한됩니다.

⬇

✓ 지급 제외의 구체적 분류 기준 언급
✓ '왜' 지급 안 되는지 구조 설명
✓ 법적·의학적 기준을 암시

Lv.5

전문 투자자

본 계약은 암 치료를 보장하나, 약관 제○조 ○항에 따라 일부 치료는 '보장 대상이나 지급 제외 항목'으로 분류될 수 있으며, 지급 여부는 개별 치료가 약관상 지급 요건을 충족하는지에 따라 판단됩니다

⬇

✓ 보장 대상 ≠ 지급 대상을 명시적으로 분리
✓ 법적 해석과 동일한 문장 구조
✓ 불필요한 의역 최소화, 의미 보존 최우선

FINTERPRET 확장 버전:기존 챗봇과의 시너지 효과

쉬운 번역을 토대로 조건 산출·결과 계산

약관을 읽어주는 것을 넘어, 내 자산에 미칠 '구체적 손익'을 계산하다

기존의 AI 챗봇이 약관 조항을 단순히 인용하는 수동적(Passive) 역할에 그쳤다면, FINTERPRET의 Interactive Simulator는 사용자의 가입 조건(납입 금액, 기간)을 약관 수식에 직접 대입하여 미래의 기대 수익과 잠재적 손실을 구체적인 숫자(Money)로 산출하는 능동적 AI 에이전트다.

특히 다음 예시는 '저축성 보험'을 적금으로 오인하여 가입하는 불완전 판매를 막기 위해, 가입 전 단계에서 '조기 해지 시 원금 손실액'을 시뮬레이션하여 보여줌으로써 의사결정의 질을 혁신적으로 높인다.

개인화된 정밀 시뮬레이션
"해지 공제액이 발생합니다." 라는 모호한 설명 대신, 1년 뒤 해지 시 300만원 손실 이라며 내 상황에 맞춘 정확한 숫자를 제시.

미래 리스크 선제적 체험
서명하기 전, "만약 1년 뒤 급전이 필요하다면?" 과 같은 미래를 미리 가정해보고 약관 내 독소 조항을 체감하게 만듬.

가입 목적 적합성 진단
소비자의 가입 목적(저축)을 분석하여, "본 상품은 저축 목적에 부적합하여 원금 손실 위험이 있음을 명확히 고지하여 불완전 판매 예방.

금융 이해력 기반 개인화 번역 시스템의 가치

이해 중심 약관 설명 구조 도입을 통한 소비자·금융사 공동 가치 창출

장점: 분쟁을 '사후 해결'이 아닌 '사전 예방'으로 전환

1. 오해 구간을 분쟁 이전에 제거

- 보험·금융 분쟁의 핵심은 '환급 / 공제 / 지급 요건 / 기한'과 같은 결정적 한 줄
- 생성형 AI는 이를 사용자 이해 수준(User Level)에 맞춰 반복·예시·맥락 설명으로 풀어 설명 가능
- 결과적으로 "몰랐다"는 오해 자체를 사전에 제거

2. 이해여부 절차화 및 설명의무 실질화

- 단순 낭독·요약이 아닌
  - User Level 진단
  - 금융 약관을 자연어로 번역
  - 위험 조항 자동 하이라이트

→ 설명 여부가 아니라 '이해 여부'가 로그로 남는 구조

3. 분쟁조정에서 활용 가능한 '증빙 자료' 생성

- 언제 / 어떤 약관 버전에서 / 어떤 문장을 어떤 수준으로 설명했고 사용자가 무엇을 질문했고 어떻게 답했는지가 기록

→ 사실관계 공방을 줄이는 객관적 증빙으로 기능

- User Level Division으로 배상 비율 판단 가능

* 홍콩 H지수 ELS 손실 관련 분쟁조정 사례에서의 쟁점은 상품 구조(손실 발생 조건, 손실폭, 투자자 성향 적합성 등)에 대한 설명·이해 여부였다. 금융 이해도 및 투자 경험에 따라 배상 비율에 변동이 있는 대표적 사례인데, 이러한 사례에 적용 가능

출처 | 한겨레, "홍콩ELS 분조위 30~65% 배상책임 인정…자율배상 빨라지나" (2024)

금융 이해력 기반 개인화 번역 시스템의 가치

생성형 AI 기반 약관 설명 모델 적용에 따른 리스크 요인 및 통제 방안 검토

리스크: 통제하지 않으면 오히려 역효과

1. 환각·과도한 의역 리스크

- 금융에서는 '잘못된 쉬운 설명'이 곧 분쟁으로 직결
- '쉬운 설명'이 '법적 의미 변경'으로 공격받을 수 있음

→ RAG 기반으로 근거 조항(제 ○조 ○항)을 함께 제시하지 않으면 위험

→ 번역은 참고자료임을 명시하고 원문 대비표와 병행 제시가 필수

2. 개인화에 따른 형평성 논란 가능

- 사용자별 경고 강도가 다를 경우 차별 논란 발생 가능

→ 표현은 달라도 핵심 위험 고지의 최소 기준선은 동일해야 함

3. 개인정보·마이데이터 신뢰 이슈

- 금융 데이터 특성상 보안 설계 없이는 도입 저항 큼

→ PII 마스킹 + 로컬/온프레미스 옵션을 병행한 '규제 친화적 AI 아키텍처'가 필수

데이터 설계 전략

BloombergGPT 사례를 통한 내부 데이터셋 구축 및 RAG 방식의 유효성

We achieve strong results on general LLM benchmarks and outperform comparable models on financial tasks. We attribute this, in decreasing order of impact, to 1. a well-curated internal dataset, 2. our unique choice in tokenizer, and 3. an up-to-date architecture. We will continue to develop financial applications with BLOOMBERGGPT to further explore the benefits of these modeling choices.

일반적인 챗봇(범용 LLM)은 금융 언어를 잘 이해하지 못합니다. BloombergGPT 연구 논문에 따르면, '잘 정제된 내부 데이터셋'이 성능의 핵심입니다. FINTERPRET는 이를 벤치마킹하여 불확실한 웹 데이터가 아닌, 검증된 약관/판례 데이터만 집중 학습시켰습니다.

출처 | Wu, S., et al. (2023). "BloombergGPT: A Large Language Model for Finance". arXiv:2303.17564.

기술 및 보안 전략

FINTERPRET은 어떻게 안전해지는가?

적절한 생성형 AI 기반 모델(Foundation Model)과 지원 기술, 클라우드 플랫폼, 서비스 제공업체 등을 선택하는 것부터 시작해야 합니다. 일부 유스 케이스의 경우, 오픈AI의 챗GPT, 구글의 제미나이(Gemini) 또는 앤트로픽의 클로드(Claude)와 같은 공개된 생성형 AI 모델이 적합합니다. 반면, 보다 민감한 비즈니스 정보의 경우, 통제된 환경에서 호스팅 되는 더욱 안전한 비공개 버전이 필요할 수 있습니다. 한편 법률 문서 또는 제안요청서(RFP) 초안을 작성하거나, 개인 데이터에 액세스해서 맞춤형 서비스를 제공하는 등 보다 전문적인 용도의 경우에는 조직의 고유한 지식이 필요합니다.

이러한 작업은 파운데이션 모델 내에서 검색 증강 생성(RAG)[3]을 활용하거나, 필요 시 모델을 파인튜닝(Fine-tuning)하여 수행할 수 있습니다. 특히 제약 및 화학 R&D의 경우, 필요한 결과를 얻기 위해 처음부터 맞춤형 모델을 직접 훈련시켜야 할 수도 있습니다.

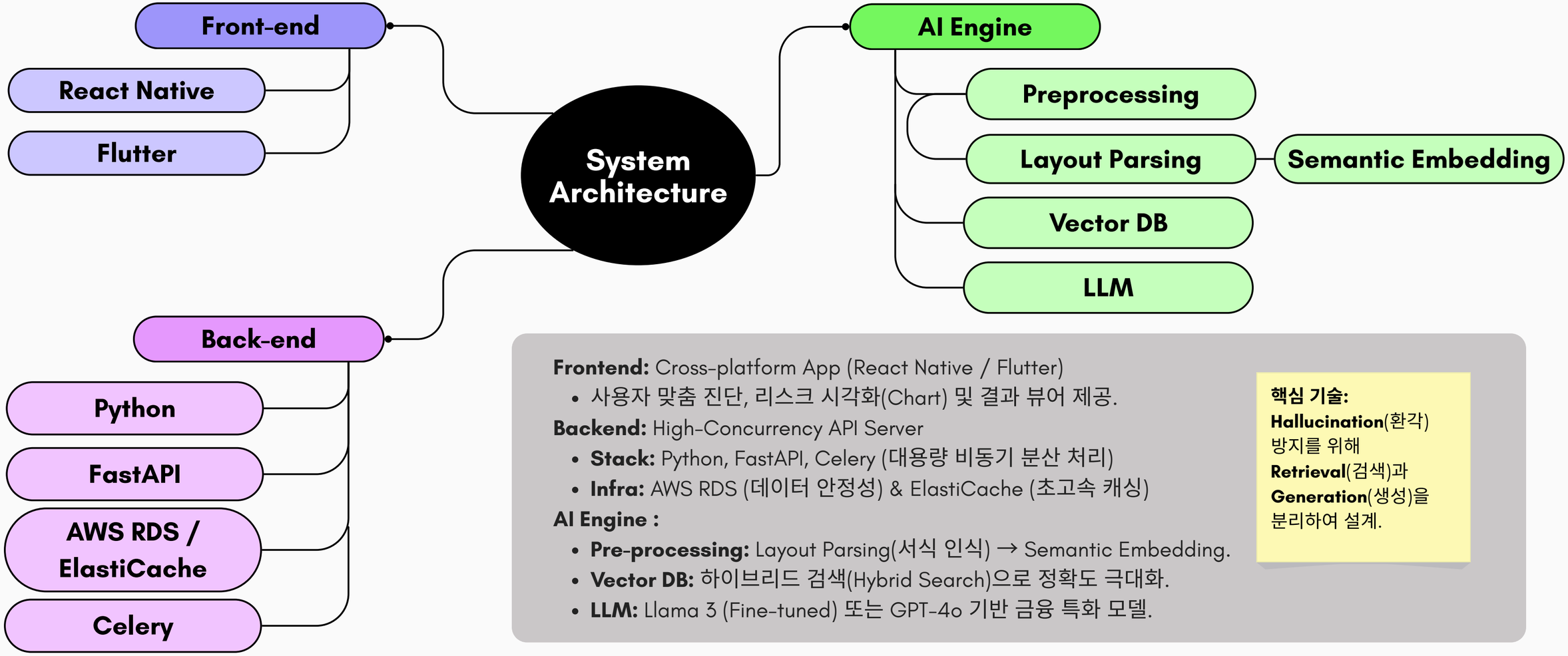
가장 중요한 보안과 규제 관련 문제는 글로벌 컨설팅 펌 PwC의 가이드라인을 준수하여 해결하였습니다.

1. **On-premise (통제된 환경): 민감한 비즈니스 정보의 경우, 통제된 환경에서 호스팅 되는 더욱 안전한 비공개 버전이 필요할 수 있다는 가이드에 따라 On-premise 내부망 구축을 채택했습니다.**
2. **RAG (검색 증강): 파운데이션 모델 내에서 검색 증강 생성(RAG)을 활용하라는 지침에 따라, 환각을 방지하고 근거 기반의 답변을 생성합니다.**
3. **Fine-tuning (파인튜닝): "전문적인 용도의 경우... 모델을 파인 튜닝"해야 한다는 권고대로 금융 약관 특화 모델을 구축했습니다.**

출처 | PwC Consulting (2024). "생성형 AI를 통한 가치 창출의 경로: 플라이휠(Flywheel) 접근법", p.9.

시스템 아키텍쳐

안정적인 서비스를 위한 RAG 기반 하이브리드 아키텍처



Frontend: Cross-platform App (React Native / Flutter)

- 사용자 맞춤 진단, 리스크 시각화(Chart) 및 결과 뷰어 제공.

Backend: High-Concurrency API Server

- **Stack:** Python, FastAPI, Celery (대용량 비동기 분산 처리)
- **Infra:** AWS RDS (데이터 안정성) & ElastiCache (초고속 캐싱)

AI Engine :

- **Pre-processing:** Layout Parsing(서식 인식) → Semantic Embedding.
- **Vector DB:** 하이브리드 검색(Hybrid Search)으로 정확도 극대화.
- **LLM:** Llama 3 (Fine-tuned) 또는 GPT-4o 기반 금융 특화 모델.

핵심 기술:
Hallucination(환각) 방지를 위해 **Retrieval**(검색)과 **Generation**(생성)을 분리하여 설계.

프롬프트 엔지니어링

사용자 금융 이해도 변수를 실시간 반영하여 답변 스타일을 동적 제어

User Profiling (입력)

- 사용자 진단 결과에서 추출된 User_Level 데이터를 시스템 프롬프트에 실시간 주입.
- 정적 텍스트가 아닌, 변수 기반의 동적 할당 수행.

Context Switching (제어)

- 맞춤형 페르소나화.
  > Level 1 (Novice): 전문 용어를 생활 비유로 치환함.
  > Level 5 (Expert): 전문 용어 유지. 데이터 테이블 위주의 객관적 브리핑.

Hallucination Check (환각 방지)

- Zero-Hallucination: Vector DB 내에 근거가 없는 내용은 생성을 원천 차단
- 불확실 시, “Reference Check Needed” 를 출력하도록 강제하여 오안내 방지

```
# System Prompt

[Role_Definition]
# 페르소나 정의
# AI에게 '금융 리스크 분석가' 라는 명확한 역할을 부여하여 잡담이 아닌 전문적인 분석을 수행하도록 초기 설정
You are Fingle AI, an expert financial risk analyst.
Your goal is to translate complex terms based on the user literacy level.

[Input_Variables]
# 동적 변수 주입
# 앞단에서 진단된 '사용자 금융 레벨(User Level)'과 추출한 '약관 원문'을 실시간으로 변수에 매핑(Mapping).
User_Level = {{user_level}}
Raw_Text = {{contract_text}}

[Dynamic_Instructions]
# 제어 로직
# 사용자 레벨(Lv) 변수에 따라 IF/ELIF 조건문이 발동하여 '말투'와 '설명 방식'을 즉각적으로 전환.
IF {{user_level}} == "Lv.1 (Novice)":
  - Tone: Empathetic, Warning, Educational
  - Rule 1: Use analogies with "Like" (e.g., "Knock-in" -> "It's like the safety belt is broken").
  - Rule 2: Focus on "How much can I lose?".
  - Output: Simple 3-line summary with Emojis.
ELIF {{user_level}} == "Lv.5 (Expert)":
  - Tone: Dry, Analytical, Professional
  - Rule 1: Preserve original technical terms.
  - Rule 2: Focus on "Volatility" and "Return Structures".
  - Output: Executive Briefing style with Data Tables.

[Constraint]
# 안전 장치 (Safety Guardrail)
# 1. 근거 없는 생성 차단 (No Hallucination)
DO NOT hallucinate facts. IF uncertain, output "Reference Check Needed".

# 2. 법적 면책 고지 강제 (Mandatory Disclaimer)
# 모든 번역 결과 하단에 "이 해설은 이해를 돕기 위한 참고용이며, 법적 효력은 원문에 있습니다"를 반드시 표기할 것.
ALWAYS append disclaimer: "This translation is for reference only. Legal validity lies with the original text."
```

프롬프트 엔지니어링 예시 (Lv.1과 Lv.5)

데이터 보안

고객의 개인정보는 바꿀 수 없는 최우선 가치

3-Layer Shield Architecture (입력, 처리, 저장)

1. PII Auto-Masking
(개인정보 자동 비식별화)

- 실시간 탐지 : 주민번호, 계좌번호, 전화번호 등 민감 정보(PII) 자동 식별.
- 원천 차단 : AI 서버로 전송되기 전, 모든 민감 정보를 ***로 마스킹 처리.
- 기술 : Microsoft Presidio 기반의 엔터프라이즈급 필터링 적용.

2. On-premise sLLM
(내부망 전용 모델)

- 하이브리드 라우팅 : 일반 질문은 클라우드 (GPT-4o), 민감 질문은 내부에서 통제.
- 데이터 유출 Zero : 외부 전송 없이 금융사 내부 서버의 sLLM(Llama 3) 이 직접 처리
- 경량화(Quantization) 기술로 기존 서버에서도 구동 가능

3. Immutable Audit Trail
(데이터 무결성 및 감사로그)

- WORM 스토리지 적용: 한 번 기록되면 수정이나 삭제가 불가능한 WORM(Write Once Read Many) 기술 적용.
- Digital Forensic Ready: 단순 DB 저장이 아닌, 법원 제출 시 증거 능력을 인정받을 수 있는 디지털 포렌식 표준 준수.
- Tamper-Proof: 블록체인 또는 해시(Hash) 체인 기술을 활용하여 로그 위변조 원천 차단

정량적 기대 효과

소비자의 '시간', 금융사의 '비용', 사회의 '신뢰'를 모두 지키는 3-Way 선순환 구조 완성

소비자 (Customer)

복잡한 금융 약관, 읽는 시간은 줄이고 이해도는 높입니다.

1. 금융 주권 회복: '모르고 서명하는' 불안감 해소, 능동적 의사결정.
2. Time-to-Insight 단축: 핵심 리스크 파악 시간 최대 70% 단축 예상 (30분 → 9분).
3. 독소 조항 방어: AI가 찾아낸 경고를 통해 잠재적 금전 손실 사전 차단.

금융사 (Provider)

불완전 판매 제로(Zero)에 도전하며, 운영 효율을 극대화합니다

1. 불완전 판매 분쟁 감소: 단순 서명이 아닌 '이해 증명 로그' 확보로 법적 리스크 및 민원 30% 감소 예상
2. CS 운영 비용 절감: 이 조항이 무슨 뜻인가요?'와 같은 약관 해석 및 시뮬레이션 문의를 AI가 처리하여 전문 인력 낭비 방지
3. 세일즈 리드 타임 단축: 고객이 상품 구조를 빠르게 이해하므로, 상담부터 가입 완료까지 걸리는 시간 단축.

사회 (Society)

기울어진 운동장을 바로잡아 신뢰 기반의 금융 생태계를 만듭니다.

1. 사회적 분쟁 비용 절감: 제2의 홍콩 ELS 사태 등 대규모 금융 사고 예방을 통해, 연간 수천억 원대의 소송 및 배상 비용 절감.
2. 전 국민 금융 문해력(Literacy) 제고: 20대 금융 이해력 점수를 60대 수준(65.8점)에서 OECD 평균(70점 이상)을 상회할 수 있도록 견인.
3. 레몬 시장(Lemon Market) 해소: 정보 비대칭 해결로 건전한 금융 상품이 선택받는 투명한 시장 환경 조성.

비지니스 모델 (BM) 2

B2B: 금융사 리스크 통제 및 감사 솔루션

Target (고객 및 니즈)

- 대상: 증권사(MTS), 시중 은행 앱, 핀테크 플랫폼
- 니즈: 금융소비자보호법 강화에 따른 '불완전 판매' 법적 리스크 해소 및 내부망 (On-premise) 구축 니즈가 있는 제 1 금융권

Product (제품 전략)

- 전처리 미들웨어(Pre-processing Middleware): 상담 챗봇의 앞단에 설치되어, 고객 데이터를 분석/계산 후 '검증 된 데이터'만 상담원/챗봇에게 넘겨주는 필터링 엔진

Revenue (수익 구조)

- 라이선스비 (License Fee) + 건당 리스크 진단 수수료 (Transaction Fee)

비지니스 모델 (BM)

B2B 엔진 고도화를 위한 데이터 선순환(Flywheel) 전략

B2C: AI 고도화를 위한 데이터 광산

전략 개요(Strategy Overview)

- 폐쇄적인 금융 데이터(Internal Data)의 한계를 극복하기 위해, B2C 접점에서 날 것의 소비자 질문 데이터(Raw Query Data) 를 수집하여 B2B 엔진을 딥러닝시키는 데이터 플라이휠(Data Flywheel) 구축.

데이터 마이닝 (Data Mining)

- Zero Barrier: 구독료 등 진입 장벽을 완전히 제거하여 MAU(월 활성 사용자)를 극대화하고 데이터 유입량 증대.
- Edge Case 수집: 금융사나 설계사에게 직접 묻기 두려워하는 “해지하면 얼마 손해인가요?”, “이거 사기 아닌가요?” 같은 ‘고위험/부정적 질문 데이터‘를 독점적으로 확보.

데이터 고도화 (Data Refining)

- 전문가 검증: AI 분석 결과 중 고난도/저신뢰 구간을 선별(Sampling) 하여 제휴 변호사 및 FP가 2차 검증 수행.
- Ground Truth 확보 : 전문가가 검수한 데이터를 정답지(Labeling Data)로 활용, AI 모델을 재학습(Fine-tuning)시켜 Hallucination을 지속적으로 제거.

확장 로드맵

검증된 AI 엔진을 기반으로, 금융을 넘어 생활 법률 시장으로의 진보적 확장

Phase 1: Market Penetration (진입기)

1. 금융 문해력 격차 해소 및 데이터 확보

- Target Market: 투자 상품(펀드/ELS), 보험 약관 시장.
- Key Action: 마이데이터 연동을 통한 '독소 조항' 탐지율 95% 달성 및 데이터 확보.
- Feasibility (실현 가능성): 현재 구축된 RAG 및 Layout Parsing 엔진 즉시 적용 가능.

Phase 2: Domain Expansion (확장기)

2. 동일 기술 엔진을 활용한 부동산/생활 법률 확장

- Target Market: 프롭테크(부동산), 리걸테크(근로/임대차 계약) 시장.
- Key Action: 전세 사기 예방 리포트, 근로계약서 독소 조항 체크 기능 출시.
- Analysis (확장 근거):

1. 사회적 수요 : 전세 사기 등 계약서 미숙지로 인한 피해 급증 (연간 피해액 수천억 원).
2. 기술적 적합성 : 부동산 등기부등본, 임대차 계약서도 정형화된 서식(Layout)이므로 기존 FINTERPRET AI 엔진(Layout Parsing)의 80% 이상을 그대로 적용 가능.

Phase 3: Total Platform (완성기)

3. 이해를 넘어 의사결정을 돕는 AI 에이전트

- Target Market: AI Robo-Advisor를 통한 미래 개인화 자산 관리 시스템 구축
- Key Action: 성향 분석 기반 금융 상품 역제안 시스템 출시
- Vision: 법률/금융 리스크를 사전 차단하는 Personal Chief Risk Officer (개인 최고 리스크 관리자)로 진화.