

# 보유 AI 솔루션

SMANU — 식품·건강기능식품 한글 표시사항 AI 워크스페이스 | 운영사: (주)파이안 | 작성일: 2026-07-07

## 솔루션 분류 정보

|        |                                                                                          |
|--------|------------------------------------------------------------------------------------------|
| 솔루션명   | SMANU — 식품 표시사항 AI 검수·초안 생성 플랫폼                                                          |
| 솔루션 목적 | 품질 보증                                                                                    |
| 솔루션 종류 | AI SaaS 솔루션                                                                              |
| 활용 분야  | 컴퓨터비전(라벨 이미지의 문자·좌표·면 인식), 자연어 처리(항목 의미 대조·법령 질의 라우팅), 생성형 AI(표시사항 초안 생성·법령 답변 정리)       |
| 적용 대상  | 식품·건강기능식품 제조/유통 기업의 QA·R&D·규제 대응(RA) 부서                                                  |
| 법령 기반  | 「식품등의 표시기준」(식약처 고시 제 2025-60 호, 현행 시행)·「건강기능식품의 표시기준」(식약처 고시 제 2025-61 호, 현행 시행)·식품위생법 등 |

## AI 솔루션 소개

### 도입 배경

식품 회사의 QA 담당자에게 한글 표시사항 검토는 가장 반복적이면서 가장 무거운 업무입니다. 품목제조보고서, 원료 배합비, 영양성분 시험성적서, 그리고 디자인 시안 — 담당자는 이 문서들을 한 줄씩 눈으로 대조합니다. 원재료 함량 표기 하나, 알레르기 유발물질 괄호 하나가 어긋나면 출고가 멈추거나 회수가 발생합니다. 「식품등의 표시기준」은 계속 개정되고, 제품 하나의 리뉴얼마다 같은 대조 작업이 처음부터 반복됩니다.

이 일을 잘하는 데 필요한 것은 더 많은 인력이 아니라, 같은 전문가가 같은 판단을 더 빠르게 내릴 수 있는 도구라는 것이 SMANU의 출발점입니다.

### 솔루션 개요

SMANU는 식품·건강기능식품의 한글 표시사항 업무를 위한 클라우드 기반 AI 워크스페이스입니다. 네 가지 기능이 하나의 작업 화면에서 연결됩니다.

- AI 검수 (Model 1) — 라벨 디자인과 품목제조보고서 등 근거 문서를 함께 읽어 차이를 항목 단위로 제시
- 초안 생성 (Model 2) — 품목제조보고서로부터 한글 표시사항 초안을 구조화하여 제안
- 디자인 비교·재검수 (Model 1 Plus) — 라벨 개정 전·후를 위치 단위로 비교, 검증된 결과 기준 재검수

• 법령 통합 검색 — 식약처·식품안전나라·법제처 등 공공 출처를 한 화면에서 조회, 출처별 원문 인용 SMANU의 원칙은 “결정은 전문가가, 가속은 AI가”입니다. AI는 담당자가 봐야 할 곳을 먼저 표시하고 검토에 걸리는 시간을 줄이는 역할을 하며, 어떤 항목도 담당자의 확인 없이 통과되지 않습니다.

## 세부 설명

아래 6개 기술 영역으로 구성됩니다. 각 영역은 실제 서비스에 적용되어 동작 중인 기능만 기재하였으며, 세부 모델 구성과 파이프라인 설계는 영업비밀로서 기술 협의 시 제공합니다.

### [1] 문서 인식 — 이중 OCR 교차 추출

- ① **입력** — 라벨 디자인(PDF·이미지), 품목제조보고서·시험성적서 등 근거 문서(PDF·이미지·Excel·Word)
- ② **처리** — 서로 다른 방식의 문서 인식 엔진 2종을 병행 운용하여 텍스트와 문자·단어 단위 좌표를 추출합니다. 칼선(다이라인) 인식으로 라벨의 면(주표시면·정보표시면 등)을 구분하고, 표시 항목을 해당 면에 매핑합니다. 저신뢰 인식 결과는 노이즈 게이트로 억제하여 오탐을 줄입니다.
- ③ **출력** — 면·항목 단위로 구조화된 텍스트 + 위치 좌표 데이터
- ④ **특장점** — 두 엔진의 강점이 달라(좌표 정밀도, 특수 단위 문자 처리 등) 교차 구성으로 상호 보완하며, 한 엔진에 장애가 발생해도 대체 경로로 처리가 이어지는 다단계 폴백 구조입니다.

### [2] 항목 대조 — 복수 AI 모델 교차 검토 (Model 1)

- ① **항목 스키마** — 표시기준 고시의 항목 구조(주표시면·정보표시면·영양정보·추가표시사항)를 따르는 필드 스키마 기반으로 추출·대조합니다.
- ② **교차 검토** — 복수 상용 AI 모델이 독립적으로 추출한 결과를 합의(consensus) 검증으로 교차 확인하여 단일 모델 오류를 줄입니다.
- ③ **의미 비교** — 단순 문자열 비교가 아닌 의미 단위 비교로, 표기 순서 변경·줄바꿈·괄호 구조 차이를 인지하여 대조합니다.
- ④ **결과 제시** — 결과는 “적합 / 검토 필요 / 주의” 3단계로 구분되고, 디자인과 근거 문서 양쪽에 근거 위치가 하이라이트로 표시됩니다. 담당자는 화면 위에서 값을 확인·수정하거나 항목을 무시·재정의할 수 있습니다.

### [3] 표시사항 초안 생성 (Model 2)

- ① **원료 트리** — 품목제조보고서의 원료 들여쓰기(계층)를 Excel 셀 서식·문서 파싱 좌표·AI 추출의 3계열 하이브리드로 인식하여 복합원재료의 하위 원료까지 트리 구조로 정규화합니다. Python(pandas) 기반 배합비율 합산 교정이 계층 인식 오류를 보정합니다.
- ② **규칙 점검** — 복합원재료 5% 미만 하위 생략 등 표시기준 규칙을 자동 점검하고, 알레르기 유발물질을 스캔하여 표기 문구를 구성합니다.

- ③ **구조화** — 영양성분 표는 셀 구조를 보존하여 추출·배치합니다. 생성된 초안의 각 항목에는 신뢰도 점수가 부여되어 담당자가 확인 우선순위를 정할 수 있습니다.
- ④ **발행·연계** — 초안은 항목별 편집을 거쳐 Word·Excel 문서로 발행되며, 이 문서는 Model 1 검수의 기준 문서로 이어서 사용할 수 있습니다.

#### [4] 디자인 비교·재검수 (Model 1 Plus)

- ① **자동 매칭** — 개정 전·후 디자인의 면(패널)을 내용 기반 최적 할당 알고리즘으로 1:1 자동 매칭합니다 — 면 순서가 바뀌어도 같은 항목끼리 대응되고, 추가·삭제된 면은 별도 표기됩니다.
- ② **위치 대조** — 문자 단위 좌표 추출·대조로 달라진 지점을 위치 단위로 제시하고, 변경·추가·삭제를 구분 표기합니다.
- ③ **재검수** — 검토가 끝난 기존 검수 결과를 기준(baseline)으로 새 디자인을 재검토할 수 있어, 리뉴얼 시 바뀐 곳에 집중합니다.
- ④ **재편집** — 내보낸 비교 결과(HTML)를 다시 불러와 주석을 추가·수정 후 재저장할 수 있습니다.

#### [5] 법령 통합 검색

- ① **출처 연동** — 법제처 국가법령정보 Open API 직접 연동으로 법령·고시·행정규칙의 정식 메타데이터를 조회하고, 식약처·식품안전나라·국민신문고는 검색 기반으로 함께 조회합니다.
- ② **질의 라우팅** — 질의 특성을 규칙 + 경량 분류기로 판별하여 일반 질의와 법령성 질의를 서로 다른 엔진으로 자동 분기합니다. 반복 질의는 응답 캐시로 처리합니다.
- ③ **원문 인용** — AI 답변에는 출처별 근거와 원문 URL 이 항상 첨부되어 담당자가 원문으로 바로 확인할 수 있습니다. 검수 작업과 분할 화면으로 병행 사용됩니다.

#### [6] 신뢰·데이터 보호 설계

- ① **학습 미사용** — 자체 학습 모델을 사용하지 않으며, 학습 목적 사용이 비활성화된 상용 AI API 만 사용합니다 — 고객 데이터가 AI 모델 학습에 쓰이지 않습니다.
- ② **접근 통제** — 계정은 운영자 승인제로 발급되며, 기능별 사용량이 계정 단위로 관리됩니다.
- ③ **운영 진단** — 처리 파이프라인의 이상 동작은 운영 진단 로그로 기록되어 장애 대응에 활용됩니다.
- ④ **근거 제시** — 라벨 검수 결과에는 근거 위치 하이라이트를, 법령 검색의 법령성 답변에는 출처별 원문 URL 을 함께 제시합니다. 법령성 질의가 아닌 일반 질의 답변은 출처 없이 제공되며, 화면에 답변 유형이 구분 표기됩니다.

#### 세부 추가 설명 (기술 스택 — 범주 수준)

- 웹: TypeScript 기반 웹 애플리케이션 — 최신 브라우저만으로 사용, 별도 설치·자체 인프라 불필요
- 클라우드: 국내(서울) 리전 서버리스 클라우드에서 서비스 호스팅·처리

- AI: 상용 AI API 조합 — 멀티모달 LLM(문서 이해·대조), 경량 분류 LLM(질의 라우팅), 검색 연동형 LLM(법령 검색)
- 문서 인식: 방식이 다른 상용 OCR·문서 파싱 엔진 2 종 교차 구성 + Excel(ExcelJS)·Word(mammoth) 네이티브 파싱
- 들어쓰기·계층 인식: Excel 셀 서식 · 문서 파싱 좌표 · AI 계층 추출의 3 계열 하이브리드로 품목제조보고서의 원료 계층 인식을 향상
- 전처리·교정: Python(pandas-numpy) 배합비율 합산 기반 원료 계층 교정, Python 이미지 기울기 보정 — 서버 브릿지로 런타임 수행
- 최적화: 프롬프트 캐싱·병렬 배치 처리로 응답 시간과 API 비용을 관리

※ 구체적인 모델·제공사 구성과 파이프라인 설계는 영업비밀로, 비밀유지 협약 하의 기술 협의 시 제공합니다.

## 활용 방식

- STEP 1. 문서 업로드** — 라벨 디자인(PDF·이미지)과 품목제조보고서 등 근거 문서를 업로드합니다.
- STEP 2. 문서 인식** — 이중 OCR·문서 파싱으로 텍스트와 위치 정보를 추출합니다.
- STEP 3. 항목 대조** — 복수 AI 모델의 교차 검토로 표시 항목 단위 대조를 수행합니다.
- STEP 4. 결과 제시** — “적합 / 검토 필요 / 주의” 구분과 근거 위치 하이라이트로 결과를 제시합니다.
- STEP 5. 담당자 확인** — 항목별로 확인·수정·재정의합니다. 담당자 확인 없이 통과되는 항목은 없습니다.
- STEP 6. 결과 내보내기** — 검토 결과를 문서(HTML·PDF·Excel 등)로 출력·보관합니다.

## 도입 효과

### 정성적 효과

- 라벨-근거문서 반복 대조 업무의 소요 시간 단축 — 담당자는 확인이 필요한 지점부터 검토 시작
- 표시사항 초안 작성을 백지가 아닌 구조가 잡힌 문서에서 시작
- 리뉴얼·고시 개정 시 변경 지점 중심의 재검토로 중복 작업 감소
- 법령·고시 근거 확인 동선 단축 — 모든 결과에 원문 출처 인용
- 검토 과정·결과의 문서화로 사내 기록·인수인계 품질 향상

### 정량적 효과

검토 소요 시간 단축률, 항목 검출 일치율 등 도입 기업별 정량 지표는 실증 과제를 통해 수요기업의 실제 데이터로 측정·산출하는 것을 원칙으로 합니다. 자체 보유 성능 지표는 아래 테스트데이터 결과서에 기재하였습니다.

# 기술 사양

## 사용 프레임워크

웹 표준 기반의 클라우드 서비스(SaaS)로, 최신 웹 브라우저만으로 사용할 수 있으며 별도 설치나 자체 전산 인프라가 필요하지 않습니다. AI 처리는 상용 AI 모델 API 와 이중 문서 인식(OCR) 엔진의 조합으로 구성되며, 자체 학습 모델은 사용하지 않습니다. 세부 구성은 기술 협의 시 제공합니다.

## 사용 아키텍처

### 데이터 흐름 (5 단계)

- ① **화면 층** — 웹 워크스페이스에서 라벨·근거 문서 업로드. 분할 화면으로 법령 조회 병행 가능
- ② **문서 인식 층** — 이중 OCR·문서 파싱으로 텍스트와 좌표 추출, 칼선 기반 면 인식·항목 매핑
- ③ **대조·생성 층** — 복수 상용 AI 모델의 교차 검토로 항목 단위 대조·초안 생성, 합의 검증으로 단일 모델 오류 억제
- ④ **근거 연동 층** — 법제처 Open API 등 공공 출처와 연결되어 결과에 원문 인용 부착
- ⑤ **결과 층** — 3 단계 구분·위치 하이라이트로 결과 제시, HTML·PDF·Excel 문서 내보내기

### 아키텍처 특징

- 무설치 — 브라우저만으로 사용하는 클라우드 SaaS, 국내(서울) 리전에서 처리
- 다단계 폴백 — 문서 인식·AI 처리 각 단계에 대체 경로를 두어 단일 엔진 장애에 대응
- 학습 미사용 — 자체 학습 모델이 없고, 학습 목적 사용이 비활성화된 상용 API 만 사용
- 비용·응답 최적화 — 프롬프트 캐싱과 병렬 배치 처리 적용
- 사용자 확정 원칙 — 모든 단계에서 최종 확정 권한은 담당자에게 있음

### 테스트데이터 결과서

- 문서 인식(문자 인식) 정확도: 1 차 엔진 98.9%(1,242/1,256 자 · 완전일치 43/48 필드), 2 차 엔진 98.6%(1,238/1,256 자 · 완전일치 41/48 필드) — 실제 유통 제품 라벨 1 종의 필드 원문(48 필드·1,256 자) 대비 문자 포함율 기준 내부 벤치마크. 최초 측정 2026-06, 동일 입력·동일 절차 재실행으로 수치 재현 확인(2026-07-07, 채점 스크립트·산출물 보유)
- 교차 구성 근거: 특수 단위 문자(μg 등) 인식은 위 벤치마크에서 1 차 엔진 우위(해당 필드 100% vs 75%)가 확인되고, 좌표 정밀도는 별도의 엔진 간 좌표 비교 분석(2026-05, 스크립트 보유)에서 확인 — 이중 구성으로 상호 보완
- 대조 파이프라인 검증: 실제 제품 케이스 2 건(라벨 디자인 + 표시사항 문서 + 정답셋) 기반 회귀 평가 체계 보유 — 평가 실행 기록(2026-04~05) 보관

- 항목 대조·초안 생성의 정량 지표(검출 일치율·소요 시간 등)와 대규모 표본 벤치마크는 실증 과제에서 수요기업 데이터 기준으로 측정 예정

※ 상세 결과서와 재현 절차는 기술 협의 시 제공합니다. 본 솔루션은 담당자의 검토를 가속하는 도구로, 컴플라이언스 적합 여부에 대한 최종 판단은 사용 기업의 담당 전문가가 수행합니다.

## 데이터 보유·수집현황

| 데이터 종류     | 데이터 규모                         | 데이터 상세내용                                                                                                                                                                                                                                                                                                                     |
|------------|--------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 법령·고시 데이터  | 공공 출처 연동                       | 「식품등의 표시기준」(고시 제 2025-60 호, 현행 시행)·「건강기능식품의 표시기준」(고시 제 2025-61 호, 현행 시행)·식품위생법 등 — 법제처 Open API 직접 연동, 식약처·식품안전나라 조회 연동. 룰셋의 조문·별지 인용은 법제처 원문과 대조 검증 완료(재검증 스크립트 보유, 2026-07)                                                                                                                                                |
| 검증용 라벨 데이터 | 내부 보유<br>(소규모)                 | 실제 유통 제품의 라벨 디자인 — OCR 정확도 벤치마크의 정답 원문·채점에 사용. 라벨·표시사항 문서 쌍의 대조 검증 케이스 별도 보유 (상세 내용은 협의 시 제공)                                                                                                                                                                                                                                |
| 산업별 표시 규칙  | 핵심 룰셋 2 종 +<br>보조 규칙 DB<br>6 종 | 일반 식품(고시 제 2025-60 호 기준)·건강기능식품(고시 제 2025-61 호 기준)의 표시 규칙을 별도 룰셋으로 분리 관리 — 검증 시 해당 산업의 규칙만 적용. 보조 규칙 DB 6 종: ①공통 표시규정(주표시면·정보표시면 필수항목, 글씨 크기, 내용량 허용오차, 소비기한 표시형식) ②식품유형별 필수항목(19 개 유형) ③세부 표시규정(제품명·원재료·영양성분·알레르기·고카페인) ④식품첨가물 명칭·용도·간략명 사전 ⑤영양 기준(1 일 영양성분 기준치·1 회 섭취참고량 66 종·허용오차 검증규칙) ⑥원료명 표준 사전 — 초안 생성과 결정론 검증 단계에 사용 |

※ 개정 추적 — 「식품등의 표시기준」 차기 개정(고시 제 2026-37 호, 2026-05-12 공포 · 2028-01-01 시행, 일부 항목은 공포일부터 시행) 사실을 인지하고 있으며, 법제처 Open API 기반 재검증 절차를 통해 개정 시행 시 룰셋과 인용 기준을 갱신합니다.

## 해당 솔루션 관련 사업실적

| 수요기업명 | 사업기간 | 과제명                                      |
|-------|------|------------------------------------------|
| -     | -    | - (신규 솔루션 — 실증 과제를 통한 첫 도입 실적 확보를 목표로 함) |

도입·실증 협의 및 데모 신청은 SMANU 홈페이지 또는 담당자 문의를 통해 가능합니다.