

전문가의 자동화

Communications Mining &
Document Understanding

강민오 / Sales Engineer



Document Understanding은 **문서**를 처리하고, Communications Mining은 **대화**를 처리합니다.

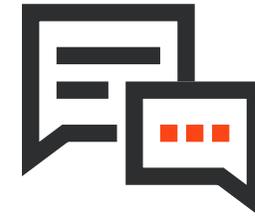
UiPath™ Document Understanding



✓ 특정한 구조가 있는 문서부터 구조가 없는 문서까지 다양한 종류의 문서를 처리

- 폼 양식의 문서
- 여권
- 인보이스
- 영수증
- 계약서

UiPath™ Communications Mining

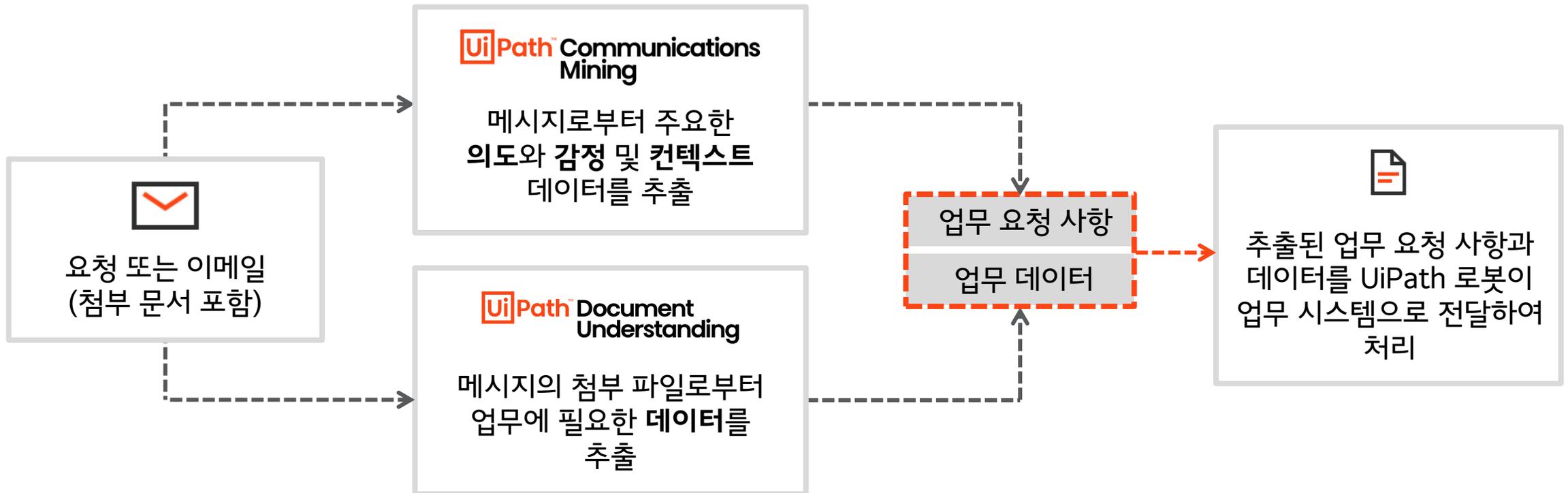


✓ 특정한 구조가 없는 텍스트 대화를 처리

- 이메일
- 서비스 티켓
- 소셜 미디어 게시글
- CRM 노트

솔루션 활용

이메일이나 티켓과 같이 요청 내용과 첨부 문서를 함께 참고해야 하는 트랜잭션 업무라면 **Communications Mining**과 **Document Understanding**을 함께 사용하여 자동화할 수 있습니다.



Communications Mining 소개



모든 업무는 메시지를 매개로 하여 이루어집니다.



고객들은 원하는
것을 얻기 위하여
연락합니다.



동료들은 협업을
위하여 대화를
합니다.



메시지는 해결해야
하는 문제를 담고
있습니다.



메시지에는 새로운
사업 기회들이
있습니다.



비즈니스는
메시지를 주고
받으면서
이루어집니다.

기존의 메시지 처리 방식으로는 효율성을 제고하기가 어렵습니다.

서베이 / 샘플 조사

- ✓ 조사 결과는 과거의 데이터를 보여주기 때문에 실시간으로 변하는 현실을 반영하지 못함.
- ✓ 비용 대비 효율성이 떨어지고 확장성이 없음.
- ✓ 이런 방식은 자동화 시스템과 연계하기가 어려움.

키워드 또는 규칙 기반의 처리

- ✓ 여러 건의 이슈와 문의를 담고 있는 메시지는 처리하기 어려움.
- ✓ 유연성이 떨어지고 작은 변화에도 키워드와 규칙을 지속적으로 변경해야 하는 비효율이 발생함.
- ✓ 메시지 내의 미묘한 특징이나 변형을 간파하기가 어려움.

웹 양식 또는 티켓 기반의 처리

- ✓ 휴먼 에러가 개입되기 쉽고 잘못된 분류가 발생할 수 있음.
- ✓ 고객의 의도하지 않은 또는 선호하지 않는 다른 채널로 메시지가 전달될 수 있음.
- ✓ 이미 알려진 이슈만 해결할 수 있고, 제한된 카테고리 내에서만 선택이 가능함.

기존의 방법들은 자동화로 연결하기 위한 구조적인 데이터를 만들어주지 못합니다.

메시지 처리를 자동화하면 지속적이고 확장 가능한 서비스가 가능합니다.

48%

단순 이메일 처리와 관련된 일에 소비되는 직원들의 시간

30%

과도한 이메일이 주의를 산만하게 만들어서 실질적인 업무에 가장 큰 방해가 된다고 생각하는 직원의 비율

78%

고객 응대 업무에서 데이터를 활용하는 보다 효율적인 방법이 필요하다고 생각하는 조직의 비율

Communications Mining을 활용하면 지속적이고 확장성이 있는 서비스가 가능합니다.



고객이 원하는 것을
정확하게 이해할 수 있다.



모든 요구사항을
실시간으로 추적하고
파악할 수 있다.



모든 요청을 자동화할 수
있다.

신속하고 확장가능한 방식으로 모든 메시지를 이해하고 처리한다.

1 연결

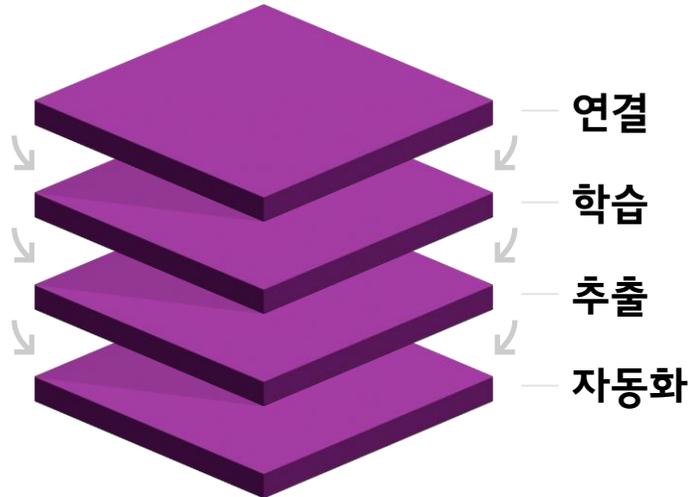
2 해석

3 자동화

- 이메일
- 티켓
- 서비스 데스크
- CRM 노트

이메일, 티켓, 노트, 서베이, 케이스, etc.

UiPath™ Communications Mining



AI NLP 기술로 실시간 메시지 해석 및 데이터 추출

- API
- ERP
- RPA**
- CMS
- CRM



구조적인 데이터가 후속 프로세스의 자동화를 가능하게 한다.

엔티티와 레이블

UiPath Communications Mining Underwriter Broker Support - Generated Small

Train Discover Explore Validation **Reports** Models Streams Settings

Dashboard **Label Summary** Trends Segments Threads Comparison

CHART LABELS

Filters Clear all

Entities ?

- + Add entity filter
- All
- Address Line
- City / Town
- Claim Date
- Country
- Email Address
- Excess Amount
- Expiry Date
- Limit Amount
- Policy Number
- Premium Amount
- UK Postcode

Labels ?

Label Treemap

The treemap visualizes the distribution of labels. The largest category is 'Policy', which is further divided into 'Claim' and 'Reason for Loss'. 'Claim' is subdivided into 'Quote' and 'Payment'. 'Reason for Loss' is subdivided into 'Urgent' and other smaller, unlabeled categories.

메시지를 해석하여

제목: **마이크로소프트 오피스** 라이선스 만료

날짜: 2024년 6월 10일 오전 8:22

발신: 홍길동

수신: it-support@company.com

안녕하세요.

저의 **마이크로소프트 오피스** 라이선스가 만료되어 지난 주부터 오피스를 사용하지 못하고 있습니다.

이번 주에 고객 대상으로 프리젠테이션이 있으니 신속하게 처리를 부탁드립니다.

홍길동 (직원번호: 1234)

인사이트를 추출

엔티티

소프트웨어:	마이크로소프트 오피스
직원번호:	1234

레이블

시스템 이슈 > 라이선스 만료

메시지 톤

긴급



No Code

가이드가 제공되는 직관적인 사용자 인터페이스로 누구나 손쉽게 사용 가능함.



State-of-the-art AI models

미리 학습된 대형 언어 모델 기반의 AI NLP 기술을 사용하여 메시지에서부터 정확한 인사이트를 도출함.



Attended & Unattended Automation

최신의 AI 기술과 연결하여 안정적으로 유인 또는 무인으로 엔드-투-엔드 자동화가 가능함.



Fully customisable

목적에 맞게 맞춤형 모델을 학습하여 메시지에서부터 의도와 테마와 감정 등을 정확하게 추출함.



Hyper-Scalable

엔터프라이즈 플랫폼으로 조직의 확장에 맞추어 탄력적으로 확장이 가능함.



Real-time monitoring and alerts

빌트인 분석 기능, 재구성 가능한 대시보드, 이메일 알림 및 리포트 기능으로 실시간 모니터링이 가능함.



AI Guardrails

데이터 보호, 규정 준수, 사람의 검증 그리고 안전한 AI 사용을 위한 보호장치가 내장되어 있음.



Easy to integrate

기존 시스템 연동 및 RPA 자동화가 손쉽게 가능하도록 API와 커넥터를 제공함.



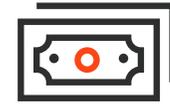
효율성, 생산성, 확장성 제고

불필요한 서비스 프로세스를 식별하여 제거하고, 대화 기반의 업무를 자동화하여 비즈니스의 생산성을 제고한다.



고객 경험의 혁신

고객이 진정으로 원하는 것이 무엇인지 파악하여 제품과 서비스를 고객의 니즈에 맞게 수정한다.



디지털 혁신의 ROI 증대

보다 신속하고 정확하게 경영 정보를 파악하고, 가장 가치 있는 변화의 기회를 식별한다.

Document Understanding

소개



1 문서 수신



문서

- 다양한 언어 지원
- 다양한 문서 타입과 포맷 지원
- 테이블/체크박스 추출
- 기울어진 스캔 문서도 처리

2 문서 이해



AI 기술을 활용한 문서 인식



Digitize



Classify



Extract



사람과 협업

추출한 내용을 사람이 검증
예외 처리

3 자동화



UiPath 자동화

- 시스템에 입력
- 이어지는 자동화 업무 연계

Built for: RPA 개발자와 시민개발자 모두 사용 가능 - 데이터 사이언스 스킬이 필요하지 않음

Built-in AI: 문서 인식 전용 AI 모델과 생성형 AI 모델을 함께 활용

Built as: SaaS, APIs, self-hosted 세 가지 형태로 deploy 가능한 UiPath Business Automation Platform

다양한 데이터 추출 방법 제공

규칙기반 추출 방법

정규표현식 추출

Form 추출

AI 기반 추출 방법

머신러닝 추출

생성형 AI 추출

추출 방법 혼합 사용 가능 (Hybrid approach)

규칙 기반의 추출 방법과 AI 기반의 추출 방법을 혼합하여 사용이 가능하다.

미리 학습된 머신러닝 모델 제공

Human resources, know your customer (KYC), employee onboarding, tax	Supply chain, shipping, manufacturing, retail	Finance, accounts payable, accounts receivable, procurement	Lending, banking	Insurance, healthcare, claims processing
Passports	Bill of Lading	Invoices Generic	Bank Statements	ACORD 25 Certificate of Liability Insurance
Utility Bills	Certificate of Incorporation/Good Standing	Invoices Japan	FM 1003 Uniform Residential Loan Application Form	ACORD 125 Commercial Insurance Application
W2 Wage and Tax Statement	Certificate of Origin	Invoices China	Financial Statements	ACORD 126 Commercial General Liability Insurance Section Form
990 Return of Organization Exempt from Income Tax - Preview	Packing Lists	Invoices India	Checks	ACORD 131 Umbrella / Excess Insurance Section Form
W9 Request for Taxpayer Identification Number and Certification	Children Product Certificate	Invoices Australia	Vehicle Titles	ACORD 140 Property Insurance Section Form
1040 US Income Tax Form	EU Declaration of Conformity	Shipping Invoices		CMS 1500 Health Insurance Claim Form
ID Cards and Driver's Licenses	Shipping Invoices	Purchase Orders		UB04
4506-T Request for Transcript of Tax Return		Receipts		
19 Employment Eligibility Verification Form		Remittance Advice		범용 문서 모델

- 모든 모델은 필요에 따라 추가로 학습이 가능하다.
- 새로운 문서에 대한 모델이 지속적으로 추가되고, 성능이 개선된다.

비즈니스 규칙 검증 (Business Rules Validation)

영수증 taxonomy

한글영수증

q 상호

w 주소

e 전화번호

r 날짜

y 총액

a 세금

g 영수증번호

h 지출구분

k 사업자등록번호

j 항목

□ 항목명

□ 수량

□ 단가

□ 항목총액

Business Rule #1

항목총액=수량*단가

Type*

Mathematical expression

Condition*

Is equal to

Math expression

수량 × * 단가 × Insert math

Business Rule #2

총액=sum(항목총액)

Type*

Mathematical expression

Condition*

Is equal to

Math expression

Sum 항목총액 × Insert math

비즈니스 규칙 설정

Details Rules ⓘ

Evaluator type ⓘ

OR ▼

Criticality level ⓘ

● Should ▼

규칙 조합 방법

OR / AND

규칙의 중요도

Should / Must

규칙의 종류

- Is Not Empty
- Mathematical Expression
- Possible Values
- Starts With
- Ends With
- Contains
- Fixed Length
- Regex
- Is Empty

규칙 예시

- 항목총액=수량*단가
- KRW / USD / EUR
- #
- MHz
- -수입-
- 10
- Wd{3}-Wd{2}-Wd{5}

신뢰도 (Confidence)

OCR 신뢰도 & 추출 신뢰도

추출 대상 항목별 신뢰도 정보

Document Type	OCR 신뢰도	추출 신뢰도
수입신고필증	N/A	
38 세번부호	98%	8541.41-9000
42 함급물량	96%	414,000 PC
54 결제금액	82%	EXW - CNY - 411,940 - TT
55 총과세가격_KRW	97%	81,714,376
관세	34%	0
부가가치세	95%	8,171,430

수입신고필증

신규번호: 4242-22-772951M | 신고일: 2022.08.01 | 세관과: 040-C1 | 일련일: 20220801 | 전자과목번호: 46

발행일: 2022.08.01 | 전자과목번호: 46

수입신고필증 (수입신고필증)

38 세번부호: 8541.41-9000

42 함급물량: 414,000 PC

54 결제금액: EXW - CNY - 411,940 - TT

55 총과세가격_KRW: 81,714,376

관세: 0

부가가치세: 8,171,430

발행번호: 2022108349838(2022.08.11) | 세관과: 040-C1 | 신고번호: 4242-22-772951M | Page: 1/2

생성형 AI를 활용한 문서 레이아웃

The screenshot displays the UiPath Document Understanding interface. The main window shows a scanned Starbucks receipt with the following text:

STARBUCKS

여의도IFC몰(L1)R점 T:1522-3232
서울 국제금융로 10
대표:송데이비드호섭 201-81-21515
[매장#3548, POS 01] 2021-11-26 10:28:17

주문번호 A-37

I-G)카페 모카	5,600	1	5,600
Lt얼음	0	1	0
L/일반휘핑	0	1	0
하트 파이	3,200	1	3,200
합계	->		8,800

결제금액 8,800

The right sidebar shows a classification tree with the following categories:

- 일반 영역 (+)
- N 상호
- V 주소
- T 총액
- D 날짜
- A 세금
- R 영수증번호
- B 사업자등록번호
- P 전화번호
- 분류 영역 (+)

생성형 AI를 활용한 문서 분류

생성형 분류기 프롬프트

문서 형식	프롬프트
영수증	영수증은 일반적으로 구매 항목, 가격, 총합계, 세금, 상점 이름 및 주소 등의 정보를 포함한다.
세금계산서	세금계산서는 일반적으로 공급자 정보, 수취자 정보, 거래 내역, 세액, 발행 날짜 등의 정보를 포함한다.
수입신고필증	수입신고필증은 일반적으로 신고인 정보, 수입물품 정보, 신고번호, 신고일자, 세금 내역 등을 포함한다.

저장 취소

생성형 시를 활용한 데이터 추출

생성형 추출기 프롬프트

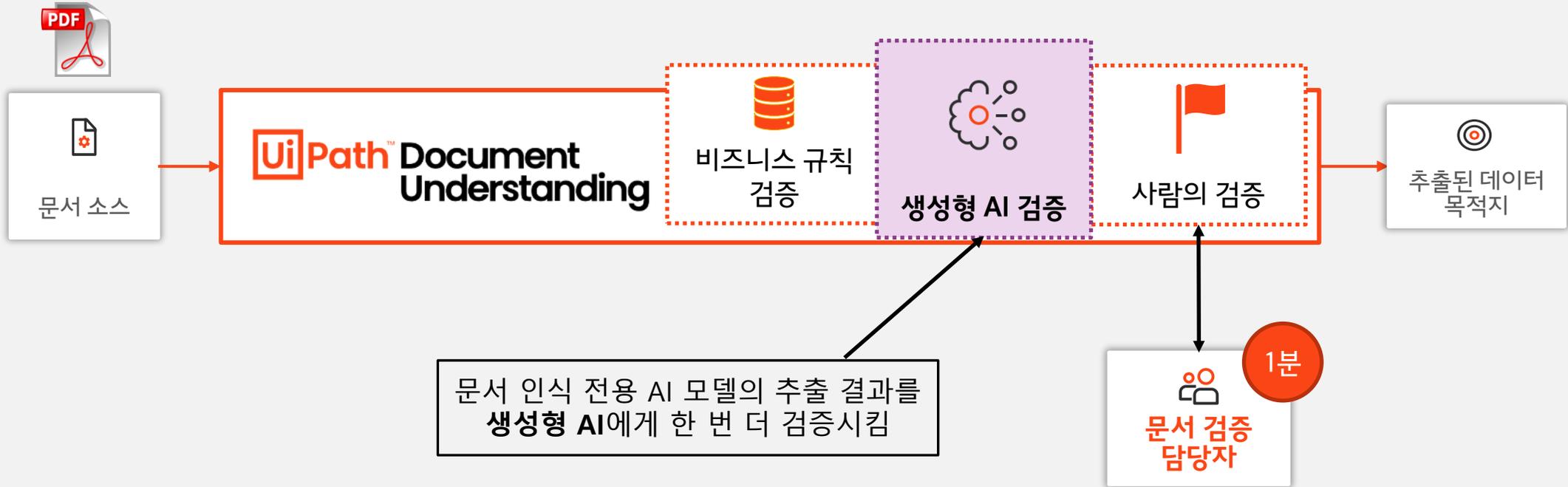
문서 형식 및 필드	프롬프트
▼ 세금계산서	
공급자-등록번호	서비스 공급자의 사업자 등록번호
공급자-상호	서비스 공급자의 상호
공급자-성명	서비스 공급자의 대표자 성명
공급자-주소	서비스 공급자의 주소
작성년월일	세금계산서를 작성한 일자
공급가액	서비스 공급가액
세액	서비스 공급가액에 대한 세금
합계금액	서비스 공급가와 세금을 합한 금액

저장 취소

생성형 시를 활용한 데이터 추가 검증



문서 처리 자동화를 통한 업무 시간 절감



100% 수작업 6 분 x 1M 문서 = 100,000 시간

50% 사람의 검증 1 분 x 0.5M = 8,333 시간
91% 업무 시간 절감

생성형 AI 검증 (40% 사람의 검증) 1 분 x 0.4M = 6,700 시간
93% 업무 시간 절감

Demo

보험사 고객 이메일 자동 처리





All Unread By Date

Monday

Prateek Gupta Policy PRC-1054 Mon 11:34 AM

Good Afternoon, I have been

Aneesh Natarajan Home and Auto Insurance Mon 10:32 AM

I would like to check out a home

Jeff Pichai Insurance - PFF953094 Mon 9:12 AM

This policy is due for renewal in 5

Andu Morie Accident and Damage durin... Mon 9:02 AM

Dear Sir/Madam, I would like to

Email Reinfer 4

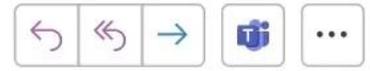
- Cancellation
- Claim
- Payment
- Quote
- Renewal

Policy PRC-1054



Prateek Gupta

To Vishal Motwani; Andu Morie; Jeremy Levin



Mon 11:34 AM

Start your reply all with:

- Ok, thank you.
- When does it expire?
- Please see attached.
- Feedback

Good Afternoon,
 I have been requested to extend the policy for an additional month.
 The policy is due to expire on the 30 June.
 We require a new policy on or before the 30 July.

Prediction: Policy>Renewal>Request



감사합니다.

