

# Document Understanding (AI-OCR/IDP 솔루션) 소개



# Table of Contents

---

1. OCR 소개
2. UiPath Document Understanding 소개
3. Demo





## 1-1 OCR 소개

OCR(Optical Character Recognition)이란 광학 문자 인식이란 뜻으로 손글씨 및 스캔된 문서를 기계가 활용할 수 있는 텍스트 데이터로 변환하는 기술입니다.

### 발전 동향

OCR 기술은 딥러닝과 기계 학습을 활용하여 정확도를 향상시키며, 자동화 응용 분야에서 더욱 발전하고 있습니다.

연도	이름	특징
1913	Fournier d'Albe	Optophone <sup>1)</sup> 개발
1929	Gustav Tauschek	Reading Machine 개발
1952	David H. Shepard	Farrington 7B <sup>2)</sup> 폰트 개발
1974	Ray Kurzweil	Omni-Font OCR <sup>3)</sup> 개발
1989	David Yang	ABBYY 설립

- 1) 시각장애인들이 사용하는 장치로, 텍스트를 스캔하고 소리로 변환해주는 장치
- 2) 현재 신용카드에도 쓰이는 OCR 인식 폰트, 숫자만 지원 가능
- 3) 글꼴에 상관없이 읽을 수 있는 OCR 소프트웨어



그림 1-2 1968년에 개발된 OCR 최적화 폰트 OCR-A



## 1-2 OCR 기술의 한계

OCR 기술은 활용도가 많지만 기업 내에서 단독으로 사용하기에는 비정형 데이터(ex: 이메일) 처리와 문서 변동성에 취약합니다. 또한, 손글씨나 난해한 서체의 인식 정확도가 낮을 뿐만 아니라 이미지 품질, 조명, 및 왜곡에 영향을 받을 수 있으며, 다국어 및 다양한 레이아웃의 문서에 대한 효과적인 대응이 부족한 한계가 있습니다.

22222		Void <input type="checkbox"/>	Employee's social security number 123-45-6789		For Official Use Only OMB No. 1545-0008	
b Employer identification number (EIN) 12-1234567			1 Wages, tips, other compensation 50,000.00		2 Federal income tax withheld 1,111.00	
c Employer's name, address, and ZIP code Readdle 795 Folsom Street San Francisco, CA 94107			3 Social security wages 35,000.00		4 Social security tax withheld 1,111.00	
d Control number R3D1			5 Medicare wages and tips 45,000.00		6 Medicare tax withheld 1,111.00	
e Employee's first name and initial Nikita V			7 Social security tips		8 Allocated tips	
Last name Bilyk			9 Verification code		10 Dependent care benefits	
123 Example Street Sample City, CA 12345			11 Nonqualified plans		12a See instructions for box 12 D 1,234.00	
f Employee's address and ZIP code			13 Statutory employee <input type="checkbox"/> Retirement plan <input checked="" type="checkbox"/> Third-party sick pay <input type="checkbox"/>		12b C 123.45	
16 State wages, tips, etc. 50,000			17 State income tax 1,535		18 Local wages, tips, etc. 50,000	
15 State Employer's state ID number OH 123-123-1234			19 Local income tax 750		20 Locality name CLEVELAND	

**Form W-2 Wage and Tax Statement 2019**  
Department of the Treasury—Internal Revenue Service  
Copy A For Social Security Administration—Send this entire page with Form W-3 to the Social Security Administration; photocopies are not acceptable.  
Cat. No. 10134D  
Do Not Cut, Fold, or Staple Forms on This Page

UIPath Document OCR - Windows 메모장  
Google Cloud Vision OCR - Windows 메모장  
파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

a Employee's social security number  
22222 Void  
123-45-6789  
b Employer identification number (EIN)  
12-1234567  
c Employer's name, address, and ZIP code  
Readdle  
795 Folsom Street  
San Francisco, CA 94107  
For Official Use Only  
OMB No. 1545-0008  
1 Wages, tips, other compensation  
50,000.00  
3 Social security wages  
35,000.00  
5 Medicare wages and tips  
45,000.00  
7 Social security tips  
2 Federal income tax withheld  
1,111.00  
4 Social security tax withheld  
1,111.00  
6 Medicare tax withheld  
1,111.00  
B Allocated tips

Ln 11, Col 22 100% Windows (CRLF) UTF-8  
Ln 1, Col 1 110% Windows (CRLF) UTF-8



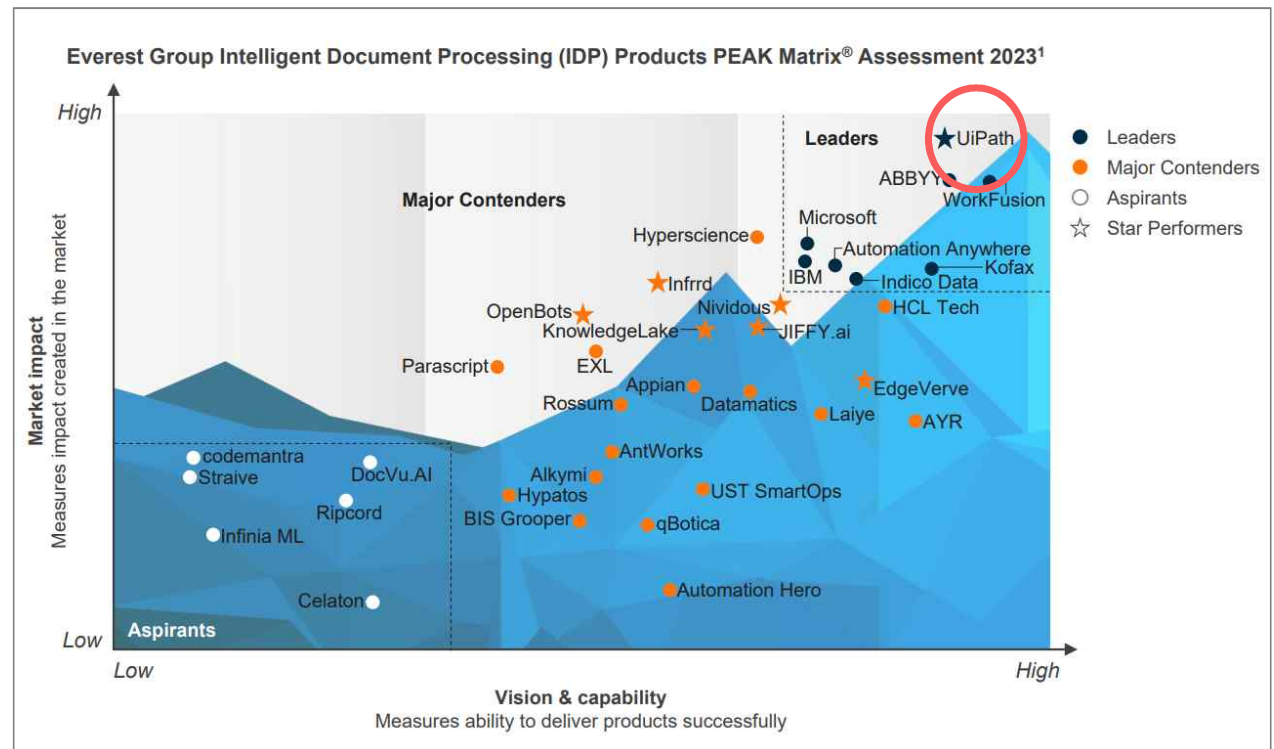
### 1-3 OCR과 IDP비교

OCR은 고정된 양식의 텍스트 인식에 주로 사용되지만, 반면에 IDP(Intelligent Document Processing)는 OCR뿐만 아니라 AI 및 다양한 기술들을 결합하여 반정형과 비정형 데이터로 이루어진 문서를 자동으로 처리하고 해석하는 기술입니다. IDP는 텍스트 추출 능력 뿐만 아니라 문서의 의미를 파악하고 분석하여 특정 작업을 수행하는 데에도 중점을 둡니다.

#### OCR – IDP 차이

	OCR	IDP
기능	이미지나 스캔된 문서에서 텍스트를 인식하여 기계가 이해할 수 있는 텍스트로 변환	OCR을 포함하여 다양한 기술을 통합하여 문서를 자동으로 처리하고 이해
작동 원리	이미지나 문서에서 문자를 텍스트로 변환하는데 주로 사용	문맥을 이해하고 의미를 추론하는 능력을 갖추어 복잡한 문서 처리에 주로 사용
기술	기계 학습과 패턴 인식과 같은 기술을 활용	OCR뿐만 아니라 기계 학습, 자연어 처리, 인공지능과 같은 다양한 기술을 종합적으로 사용
사용 사례	신용 카드 인식, 운전면허증 인식 등에 활용	계약서 분석, 금융 보고서 처리, 고객 지원 문의 처리 등에 활용

#### IDP(지능형 문서 처리)







## 2-1 UiPath Document Understanding 소개 및 기대 효과

UiPath의 Document Understanding(DU)은 AI 모델을 활용하여 다양한 업무 문서를 처리하는 기능으로 100개 이상의 언어를 지원하며, 다양한 문서 구조, 필기, 체크박스 등을 인식하고 여러 파일 형식에 대응합니다. 이를 통해 로봇은 데이터를 읽고 추출하며, 학습을 통해 생산성을 향상시키고 문서 작업에 소요되는 시간과 비용을 절감할 수 있습니다.



일반적으로 문서 내 데이터를 수동으로 처리하고 해석함



문서 처리를 위해 로봇에게 위임



다양한 유형의 문서들과 화질이 낮은 이미지를 처리하기 어려움



DU 자동 처리 및 사람의 검증



비용 증가 및 시간 소모



비용 감소 및 시간 효율성 증대



실수가 발생하면, 재작업을 하기에 손실이 발생함



AI를 통한 정확하고 빠른 문서 처리



모든 고객들의 요구를 충족시켜 줄 공급업체를 찾기 어려움

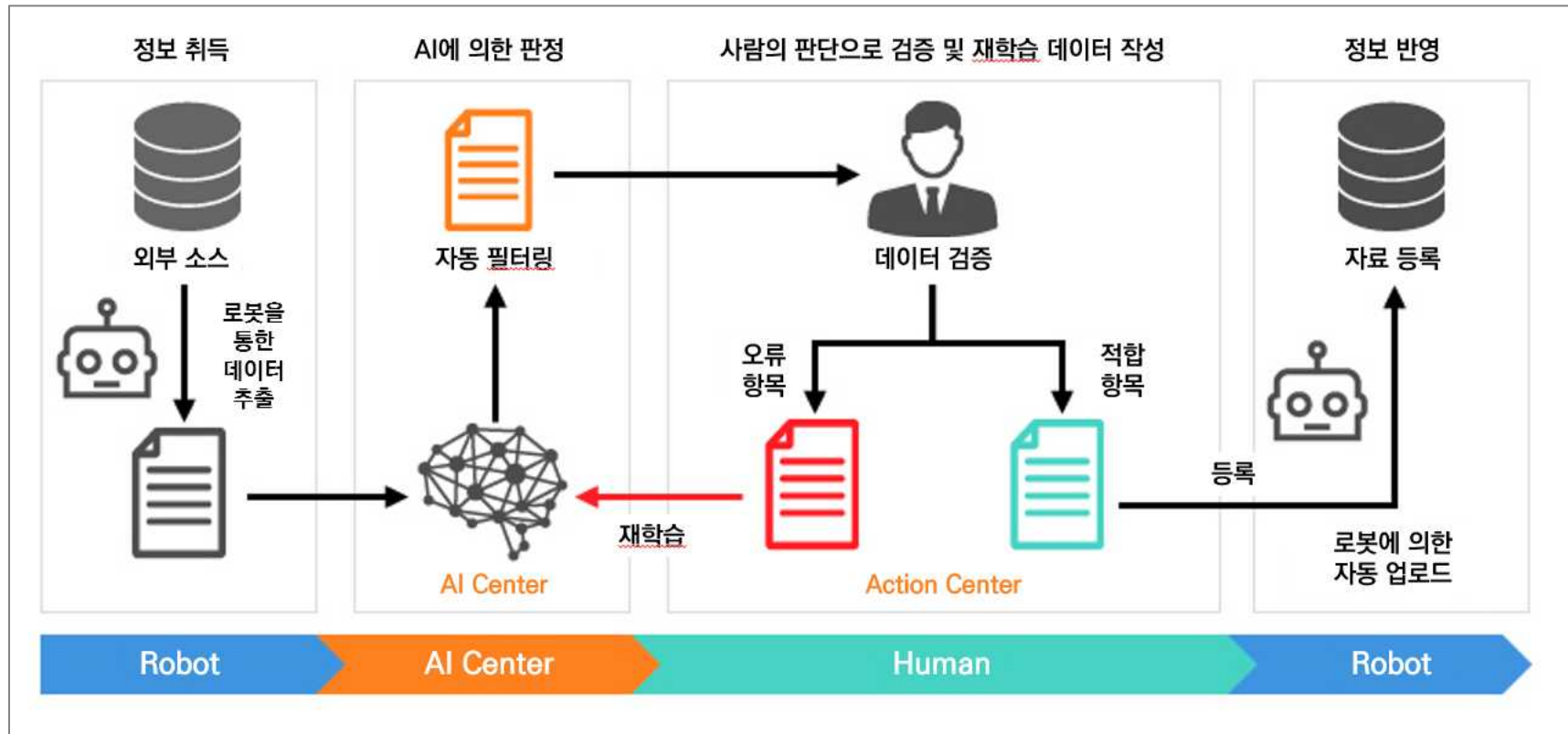


AI를 활용한 End-to-End 솔루션



## 2-2 Document Understanding 진행 과정

UiPath의 다양한 솔루션은 Document Understanding 기능을 강화하고 확장하여 더욱 효과적인 문서 처리를 가능하게 합니다. 사용자는 DU 내에서 AI 모델을 선택하여 문서를 이해하고 분석할 수 있으며, 예외 상황 발생 시 직접 검증을 통해 AI 모델을 재학습시켜 정확성과 효율성을 더욱 향상시킬 수 있습니다.





## 2-3 Document Understanding 특징1 – 레이블링(Labeling)

Labeling은 사용자가 문서의 중요한 부분을 레이블로 지정하여 AI 모델에 학습 데이터를 제공하는 핵심 작업입니다. 이를 통해 AI 모델은 다양한 문서 유형에서 필요한 정보를 정확하게 추출하고, UiPath Studio와의 통합을 통해 효율적인 작업 흐름을 갖출 수 있습니다.

The screenshot displays a document understanding interface. The main area shows a W-2 form with various fields labeled with colored boxes and text. The labels include:

- 22222 (Employer identification number)
- 123-45-6789 (Employee's social security number)
- 12-1234567 (Employer identification number (EIN))
- 35,000.00 (Social security wages)
- 45,000.00 (Medicare wages and tips)
- 50,000.00 (Wages, tips, other compensation)
- 1,111.00 (Federal income tax withheld)
- 1,111.00 (Social security tax withheld)
- 1,111.00 (Medicare tax withheld)
- 123.45 (Allocated tips)
- 1,234.00 (Dependent care benefits)
- 123.45 (Nonqualified plans)
- 123.45 (See instructions for box 12)
- 123.45 (Statutory employee)
- 123.45 (Retirement plan)
- 123.45 (Third-party sick pay)
- 123.45 (Other)
- 123.45 (State wages, tips, etc.)
- 1,535 (State income tax)
- 50,000 (Local wages, tips, etc.)
- 750 (Local income tax)
- CLEVELAND (Locality name)

The sidebar on the right lists the fields and their corresponding values, with checkboxes indicating whether they are labeled:

- SSN: 123-45-6789
- Employer FED ID Number: 12-1234567
- Employer Address: Readle 795 Folsom Street San Francisco, CA 94107
- Employee Name: Void
- Employee Address: 123 Example Street Sample City, CA 12345
- 1 Wages: 50000.00
- 2 Federal Income Tax Withheld: 1111.00
- 3 Social Security Wages: 35000.00
- 4 Social Security Tax Withheld: 1111.00
- 5 Medicare Wages: 45000.00

The bottom of the form shows the title "Form W-2 Wage and Tax Statement 2019" and the Department of the Treasury—Internal Revenue Service logo.





## 2-4 Document Understanding 특징2 – 예외 상황 처리

DU의 예외 상황 처리는 Action Center 및 Forms를 통해 이루어집니다. 작업 실행 중에 예외 상황이 발생하면, 해당 예외에 대해 사람이 직접 검증을 수행할 수 있습니다. **사용자는 화면에서 문서를 직접 검증하고, 이 결과는 AI 모델로 다시 전달되어 모델을 재학습시킵니다.** 이를 통해 모델은 시간이 지남에 따라 정밀도를 향상시키고, 문서 이해의 정확성과 효율성을 높일 수 있습니다.

UiPath Forms 검증

청구서 요약

KCC I&C 청구서 총 금액은 2362635.00 입니다.

청구서 처리 액션

☒ 승인  
☐ 반려

제출

60일 남음

지금 구입하기

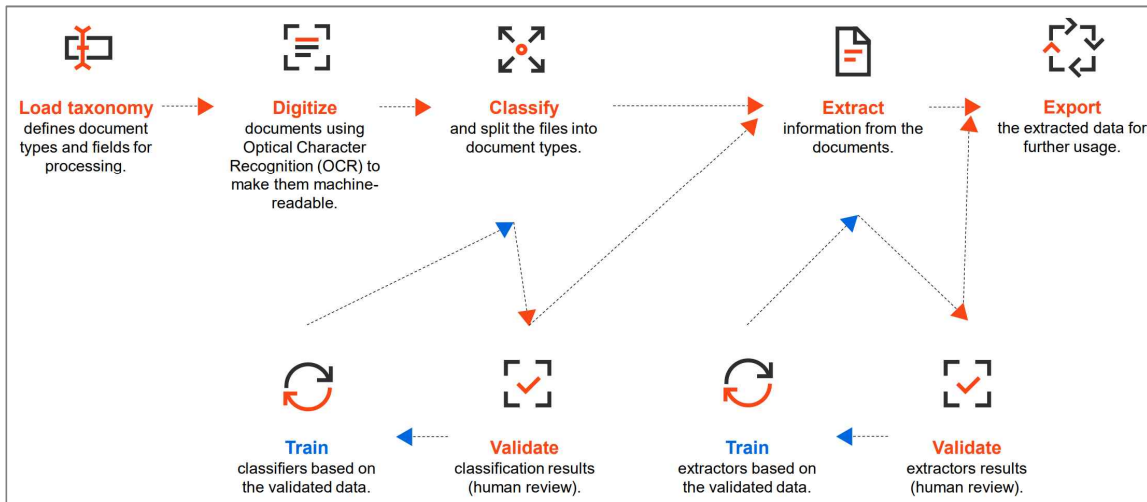
KC



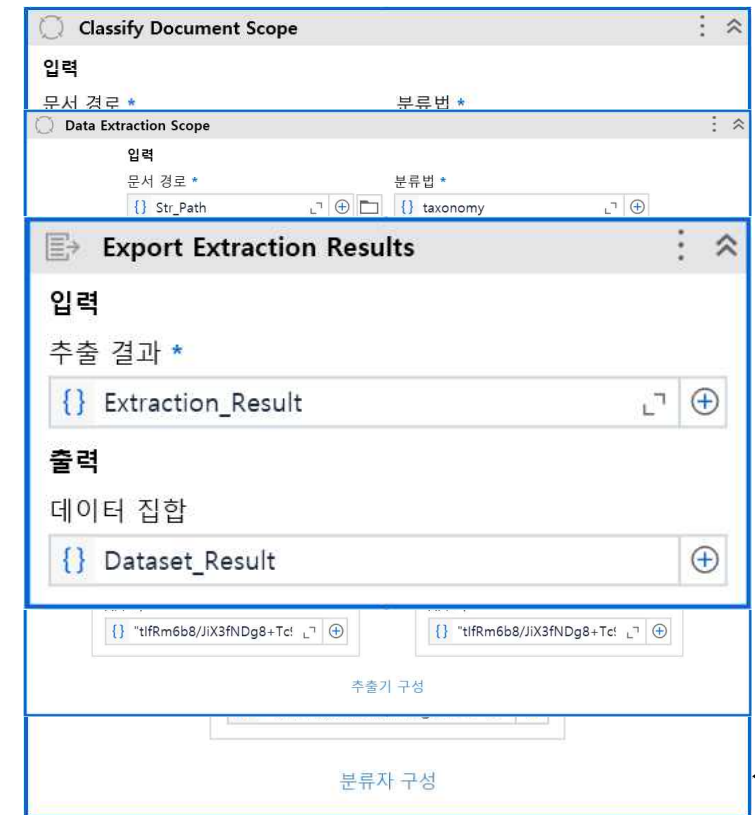
## 2-5 Document Understanding 특징3 – 쉬운 자동화

DU 프레임워크는 변화하는 환경에 빠르게 대응할 수 있는 유연한 구조를 가지고 있습니다. 정형화된 프로세스보다는 더 민첩한 방식으로 업무를 수행하며, 필요에 따라 리소스를 적절하게 조정할 수 있습니다.

### 자동화 순서



### UiPath Studio



신



## 2-6 Document Understanding 사용 분야

 <b>일반</b>	 <b>금융 서비스 및 보험</b>	 <b>인사 (HR)</b>
<ul style="list-style-type: none"><li>▪ 송장</li><li>▪ 영수증</li><li>▪ 구매 주문서</li><li>▪ 유틸리티 청구서</li><li>▪ 선적증권</li><li>▪ 여권</li><li>▪ 라이선스</li></ul>	<ul style="list-style-type: none"><li>▪ 지급 계정</li><li>▪ 매출 채권</li><li>▪ 청구 처리</li><li>▪ 대출 신청</li><li>▪ 모기지 처리</li><li>▪ 계좌 개설 및 고객 온보딩</li><li>▪ IRS 양식</li><li>▪ 공급업체 온보딩</li><li>▪ 준수 관련 프로세스</li></ul>	<ul style="list-style-type: none"><li>▪ 직원 온보딩</li><li>▪ 이력서 심사</li><li>▪ 인사 기록 관리</li></ul>
 <b>제조</b>	 <b>공공 부문</b>	 <b>의료</b>
<ul style="list-style-type: none"><li>▪ 판매 주문 처리</li><li>▪ 고객 부품 요청</li><li>▪ 송금 처리</li></ul>	<ul style="list-style-type: none"><li>▪ 이민 신청</li><li>▪ 학교 신청</li><li>▪ 여권 관리 신청</li></ul>	<ul style="list-style-type: none"><li>▪ 의료 양식</li><li>▪ 의료 청구서</li><li>▪ 건강 기록</li><li>▪ 의약품 처방</li></ul>



### 3-1 Demo1 – Invoice 데이터 추출

1 Showcase Description	✓ 메일로 전달 받은 Invoice PDF 파일에서 필요한 Data를 추출 후 주어진 양식에 입력한 다음 담당자에게 완료 메일을 발송하는 프로세스입니다.		
2 Business Categories	✓ 구매	3 Robot Type	✓ UnAttended
4 UiPath Technologies	✓ Document Understanding ✓ AI Center ✓ Apps, Trigger	5 Technologies	✓ PDF, Email



### 3-1 Demo1 – Invoice 데이터 추출

---

Invoice 샘플 메일 발송





### 3-2 Demo2 – DU + 생성형 AI

문서가 많아질수록 사람이 수작업으로 Labeling 작업을 하기에 한계가 찾아옵니다. 그럴때, 생성형 AI와 함께 작업을 하면 더 편리하게 진행할 수 있습니다.

**가져오기 상태**

가져오기 형식 원시  
OCR UiPath Chinese-Japanese-Korean OCR <https://du.uiopath.com/cjk-ocr>  
일괄 처리 receipts

가져오기 상태 20/20 100%

문서 처리됨 20  
중복 항목 건너뛴 0  
오류 0

로그 보기 닫기

**Receipt Document Content:**

KCC I&C  
용산구 한강대로98길  
서울특별시  
South Korea

Bill To:  
UiPath  
1 Vanderbilt Ave 60t  
New York  
U.S.A

#	Item Description	Qty	Rate	TAX	Amount
1	UiPath Studio Textbook	200	50	1000.00 10	10000.00

**Form Fields:**

- Name Vendor
- Vendor Address
- Phone Number
- Date
- Tax Amount
- Total Amount
- Receipt Number

**Table:**

#	Item Description	Qty	Rate	TAX	Amount
1	UiPath Studio Textbook	200	50	1000.00 10	10000.00



### 3-3 Demo3 – 능동적인 DU

생성형 AI를 사용해도 문서 Dataset에 전체 적용을 하려면 시간이 오래 걸렸지만, 능동적으로 DU가 Labeling 작업을 도와준다면 더 빠르게 진행할 수 있습니다.

UiPath Document Understanding

REST API

+

프로젝트 만들기

자체 사용자 지정 모델을 학습하기 위한 Document Understanding 프로젝트 만들기

자동화 생성

즉시 사용 가능한 모델 중 하나를 사용하여 Document Understanding 자동화 생성

프로젝트

검색

열

프로젝트 만들기

이름	문서 형식	추출기	문서 처리됨	생성 날짜	
DU_Test_DX	<a href="#">미리 보기</a> invoices	0	0	2024. 2. 6. AM 10:31:36	
DU_Test_GR	Invoices, Receipts	2	0	2024. 2. 5. AM 11:15:12	
DUTest_hjchoi	<a href="#">미리 보기</a> invoices, receipts	0	0	2024. 1. 31. AM 9:22:57	
DUTest_grcha	PO, Receipt	3	10	2023. 11. 23. PM 4:11:29	
Receipts	Receipts	0	0	2023. 11. 16. AM 10:22:44	

1 - 5/5

1/1페이지

페이지당 항목 10

감사합니다.

