



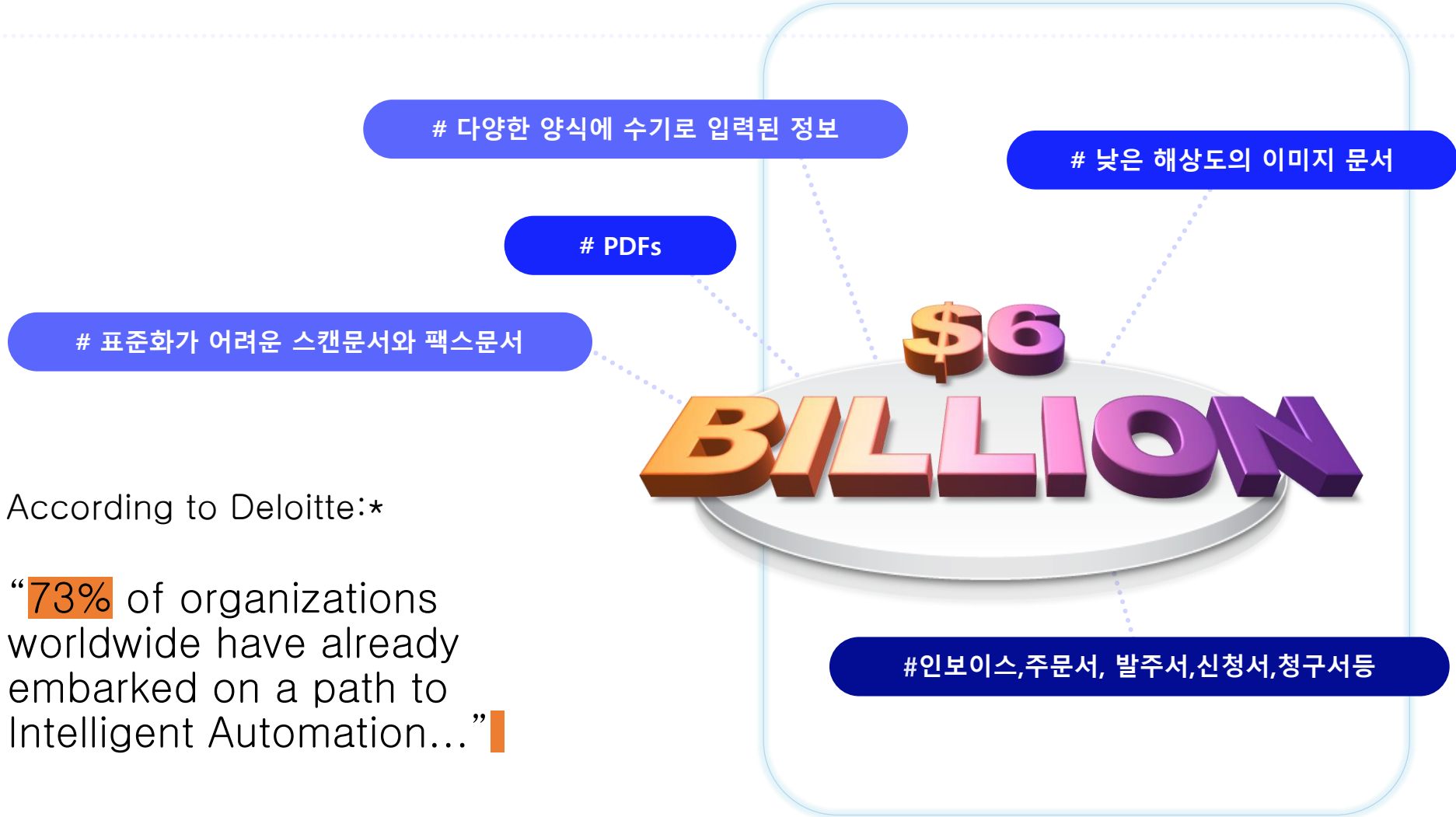
INTRO TO HYPERSCIENCE

지능형 문서 처리(IDP-Intelligent Document Processing)

 누리어시스템

현황

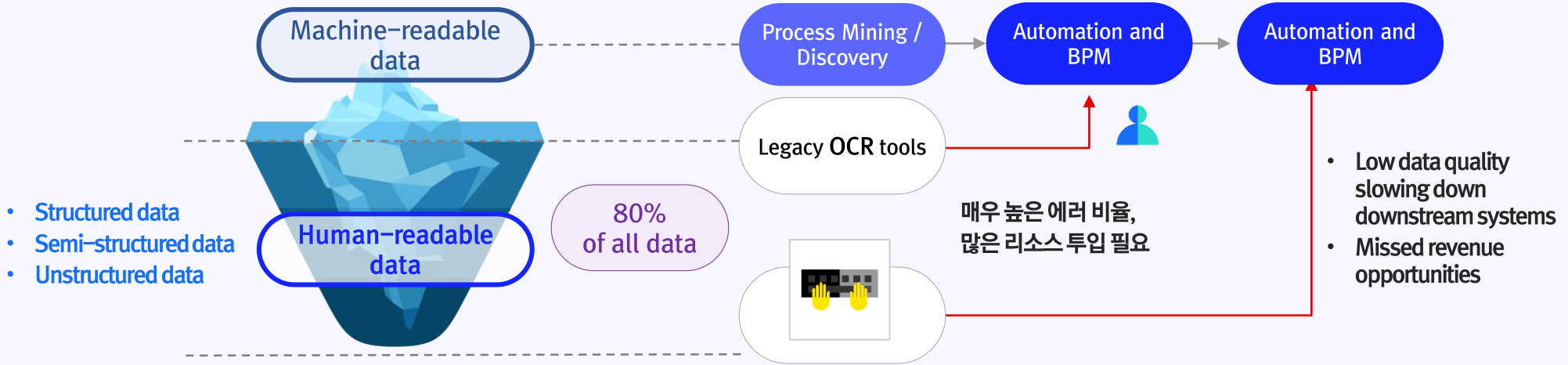
기업은 고객 및 기타 조직에서 생성된 데이터를 추출하고 사용하기 위해 매년 600 억 달러 이상을 지출하고 있습니다.



According to Deloitte:*

“73% of organizations worldwide have already embarked on a path to Intelligent Automation...”

현실의 한계



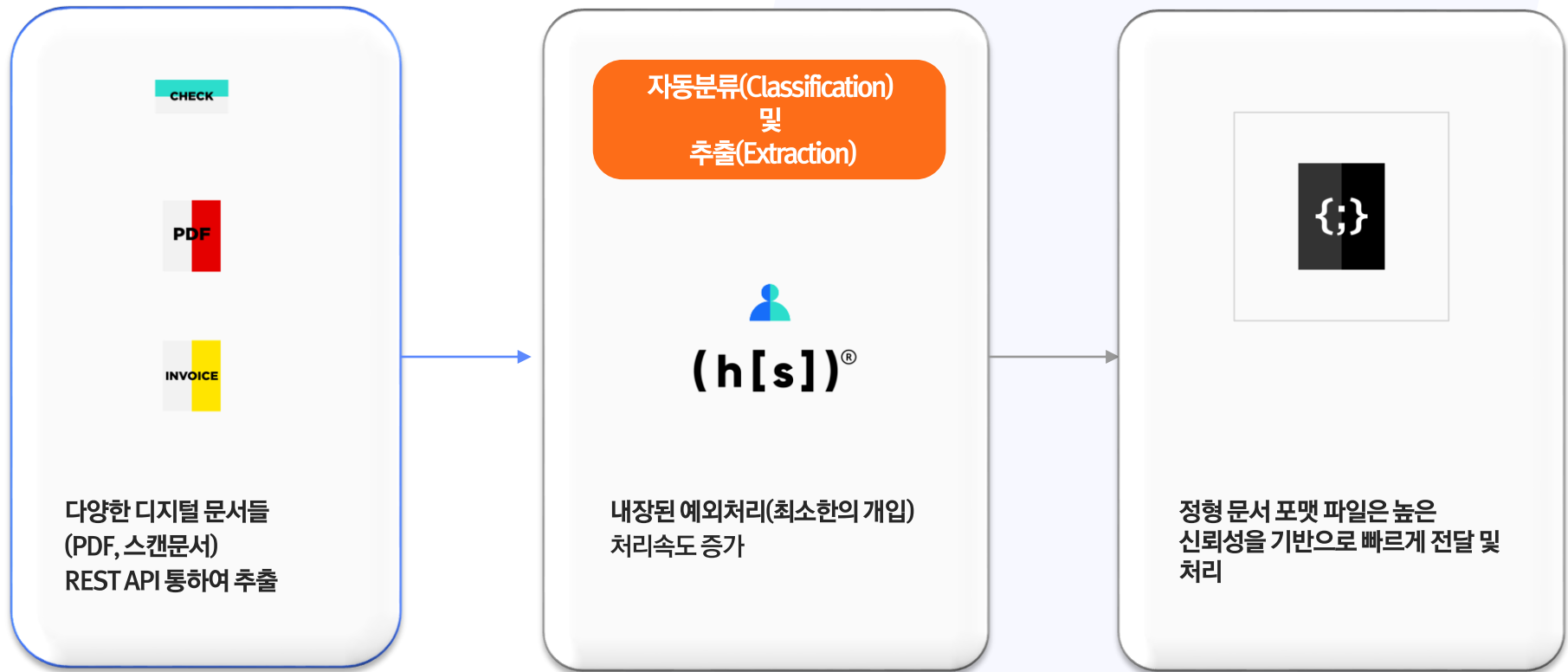
문제 #1 수동 입력 한계

- 높은 운영 비용
- 리스크 증가
- 비효율적인 인력 활용
- 낮은 고객 만족도
- 기회 비용 손실

문제 #2 OCR 한계

- 수기 인식 불가
- 고해상도 이미지 필요
- 모호한 문서 내용
- 낮은 신뢰성, 수동 검토 및 검증 필수
- 성능 개선 한계










하이퍼사이언스 솔루션



리소스 투입을 줄이면서, 높은 정확성 제공하고 처리 속도를 획기적 개선

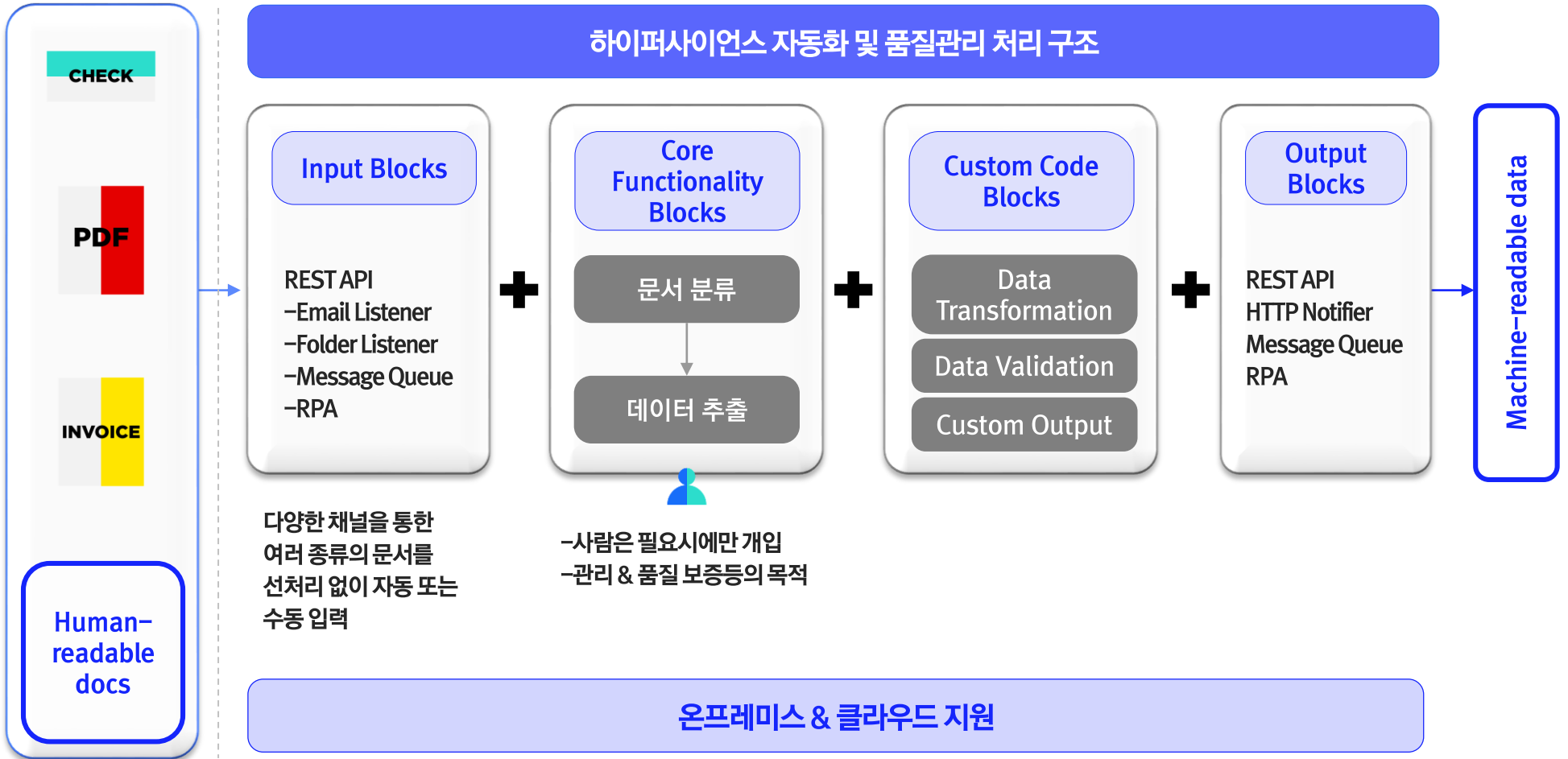
처리 플로우

리소스 투입을 줄이면서, 높은 정확성 제공하고 처리 속도를 획기적 개선

	CLASSIFICATION	EXTRACTION	HUMAN-IN-THE-LOOP EXCEPTION-HANDLING	VALIDATION	ENRICHMENT	
 HUMAN-READABLE DOCUMENTS						 MACHINE-READ ABLE DATA
INPUT	● ●	● ●	● ●	● ●	● ●	OUTCOME
 다양한 디지털 문서들 (PDF, 스캔문서) RESTAPI 통하여 추출			 약 1% 미만의 필드에 교정 개입	 DB연결, 데이터검증등 비즈니스 로직 적용	 JSON 포맷 데이터 제공	

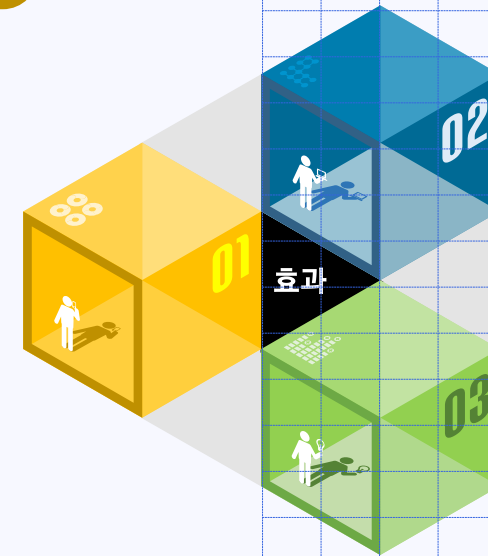
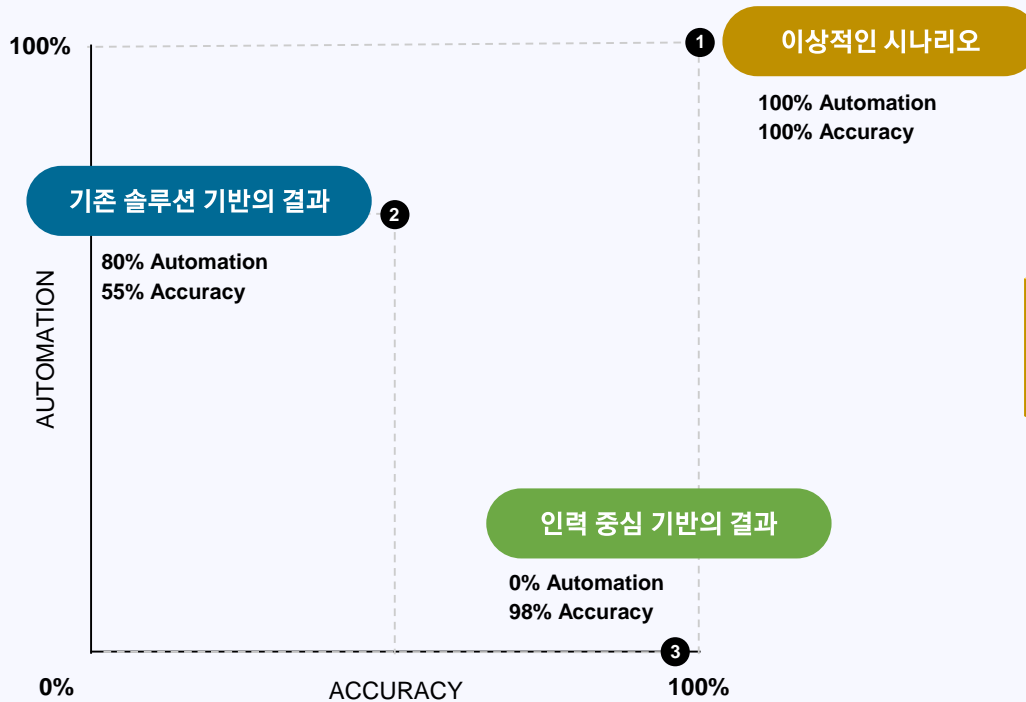
작동 구조

하이퍼사이언스 자동화 및 품질관리 처리 구조



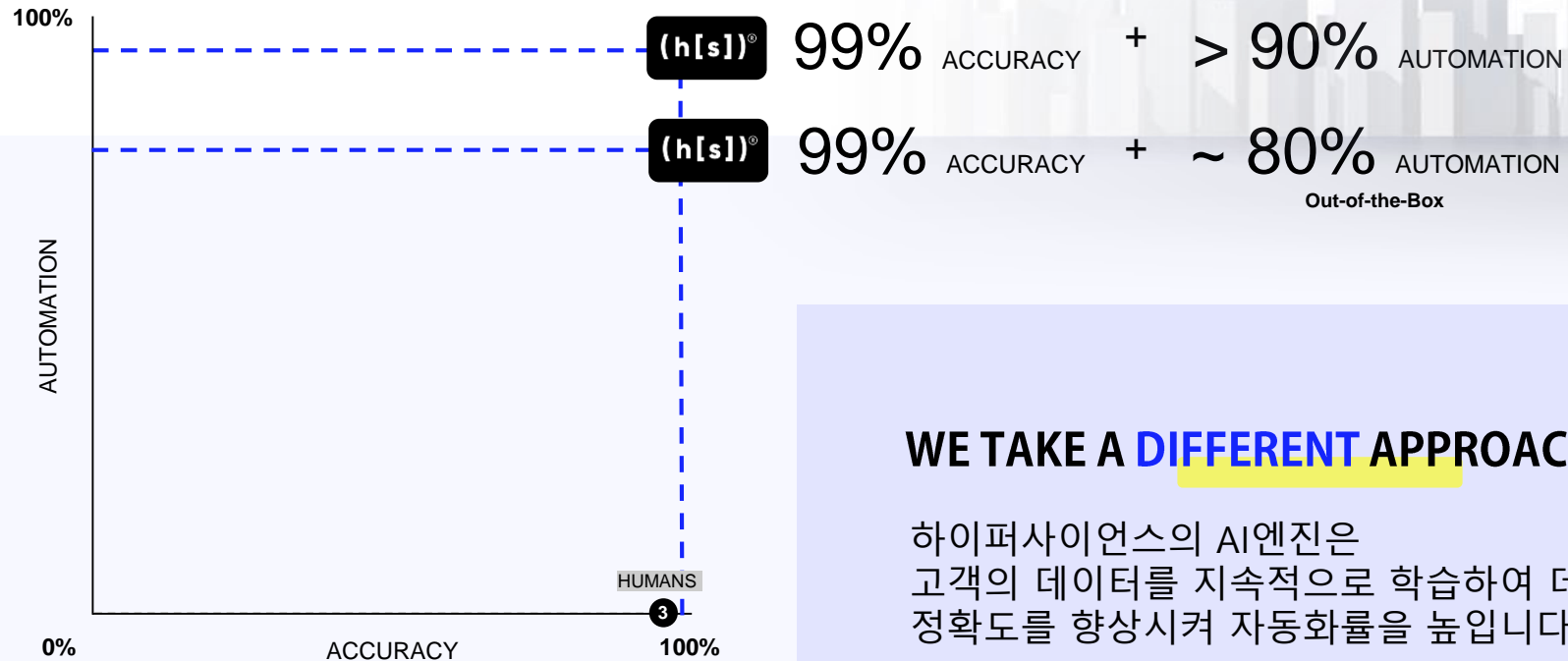
기대 효과 - 도입 전

THE AUTOMATION VS. ACCURACY TRADEOFF



기대 효과 - 도입 후

HIGHER AUTOMATION OVER TIME.



WE TAKE A **DIFFERENT** APPROACH.

하이퍼사이언스의 AI엔진은
고객의 데이터를 지속적으로 학습하여 데이터의
정확도를 향상시켜 자동화률을 높입니다.



실제 비교 검증 사례

사례 1

Name of organization (<input type="checkbox"/> Check box if name changed and see instructions) INSTITUTE FOR FAMILY-CENTERED CARE, INC.	D Employer identification number (Employer's trust, see instructions.) 52-1777133
Number, street, and room or suite no. If a P.O. box, see instructions. 6017 ARLINGTON ROAD, 309	E Unrelated business activity code

Hyperscience:

- INSTITUTE FOR FAMILY-CENTERED CARE, INC.
- 52-1777133
- 6017 ARLINGTON ROAD, 309

Microsoft OCR:

- FOR FAMV,YCENTEREO CARE, INC
- *Blank*
- RLEtctON ROAD.

OmniPage OCR:

- FOR FAMILYCENTERED CA
- 52-1-777133
- MI ARLII4G7ON ROAD, 30g

Tesseract:

- [Mmm" 'Té'ffihQEBfi'Qbéifihthfi'fiMfi'im'm I S'é'Q
- fidfim amm \$ngEgg5233115335fiéaggjzfigéfic- W -- Z
- {:143? \$ Z : 73"" @1.?'&BLE!€§19§.BOEQ%9§_

Google Cloud Vision:

- INSTITUTE FOA FAMILY-CENTEAD CARE, INC .
- 52-1777133
- 0017 ARLINGTON ROAD, 309

실제 비교 검증 사례

사례 2

Why Hyperscience? *Handwriting and signatures*

Firm's name ▶ *Williamson Preparation Services, LLC*

Hyperscience:

WILLIAMSON PREPARATION SERVICES, LLC.

Microsoft OCR:

Z..eS, Z-2-*d l

OmniPage OCR:

lr tikj:"Z(7etiet5 Oz.", `C__RS

Tesseract:

Matfwtceééc;_l

Google Cloud Vision:

Blank

실제 비교 검증 사례

사례 3

Why Hyperscience? Cross outs and corrections

16	Generation-skipping transfer taxes (from Schedule D, Part 3, col. G, total)	16	40.00
17	Total tax. Add lines 15 and 16	17	125 135.00

Hyperscience:

- 40.00
- 125

Microsoft OCR:

- I Z.,-/öYDO
- Blank

OmniPage OCR:

- 4/6 wa-a
- Mrsiittfiegg

Tesseract:

- fi 45333?)
- 1riIZZS'W'

Google Cloud Vision:

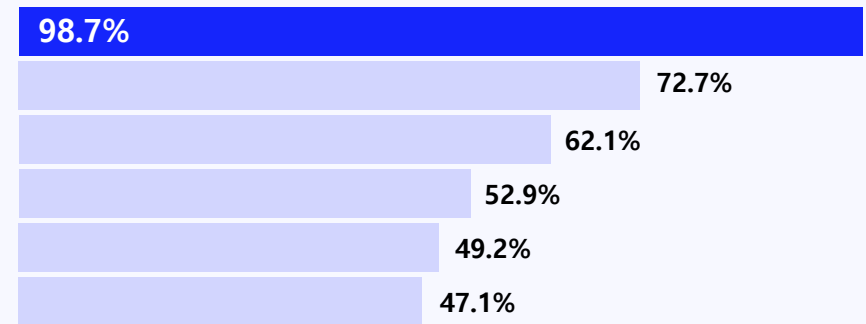
- 40.00
- 12503

검증된 수치

자동화 비율*



정확성†



Our tech outperforms competitors,

delivering the highest levels of accuracy and automation.

* Automation rate for 99.5% accuracy SLA and assuming humans are 100% accurate. Competitors' actual automation rates will be worse in production.

† Machine-only benchmarks, assuming no human-in-the-loop.

고객의 목소리

[Better] [Faster] [Cheaper]

67% 정확성 증가

5배 이상의 처리 속도 증가

80% 비용 절감



*"We got a bunch of people for this job because we thought we'd need them, but we've had to **reassign those people to unrelated work.**"*

[Fortune 500 Financial Services Firm]



*"Hyperscience is **the magic behind the curtain** here."*

[Financial Services Firm]



*"I'm getting **near 100% extraction** at really high accuracy. In the 5 years I've been here, this is the **first time I haven't heard complaints about extraction.**"*

[Global Insurer]



*"Before Hyperscience, one data-keyer was transcribing 3 forms per day... after 2 months in production they're doing **25 per day!**"*

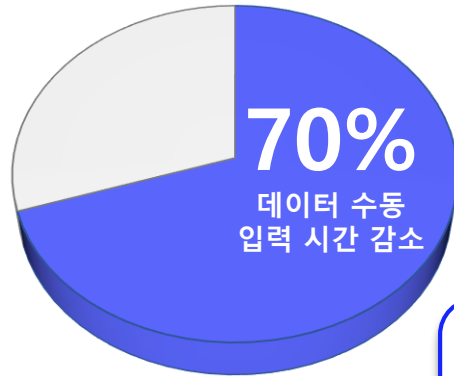
[Legal Services Firm]

도입 사례

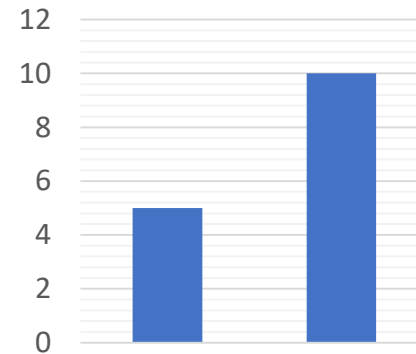
83.3%

인보이스 처리 시간 절감

[Top 10 보험 회사]



[For one insurance
company processing
16M handwritten life
insurance forms/year]



[Large financial services
firm processing account
opening applications]

3-5배 ROI,

4-6개월 이내에
투자 비용 회수



주요 차별화 요소



1 2 3 4

Rapid Time-to-Value

비교할 수 없는 높은 정확도와 자동화 비율은 수많은 레퍼런스에서 검증되었습니다.

대체 솔루션을 사용하는 것보다 훨씬 짧은 시간 내에 투자비용을 회수할 수 있습니다.

Built for Business Users

코드 없는 플랫폼은 직관적이고 사용하기 쉽게 설계되었습니다.

조직의 모든 사용자들이 쉽게 사용할 수 있어 전반적인 워크플로우의 개선을 가능하게 합니다.

Deploys With Your Infrastructure

불과 몇시간이내에 PII 문제 없이 운영될 수 있습니다.

온 프레미스 또는 프라이빗 클라우드 인스턴스를 지원합니다.

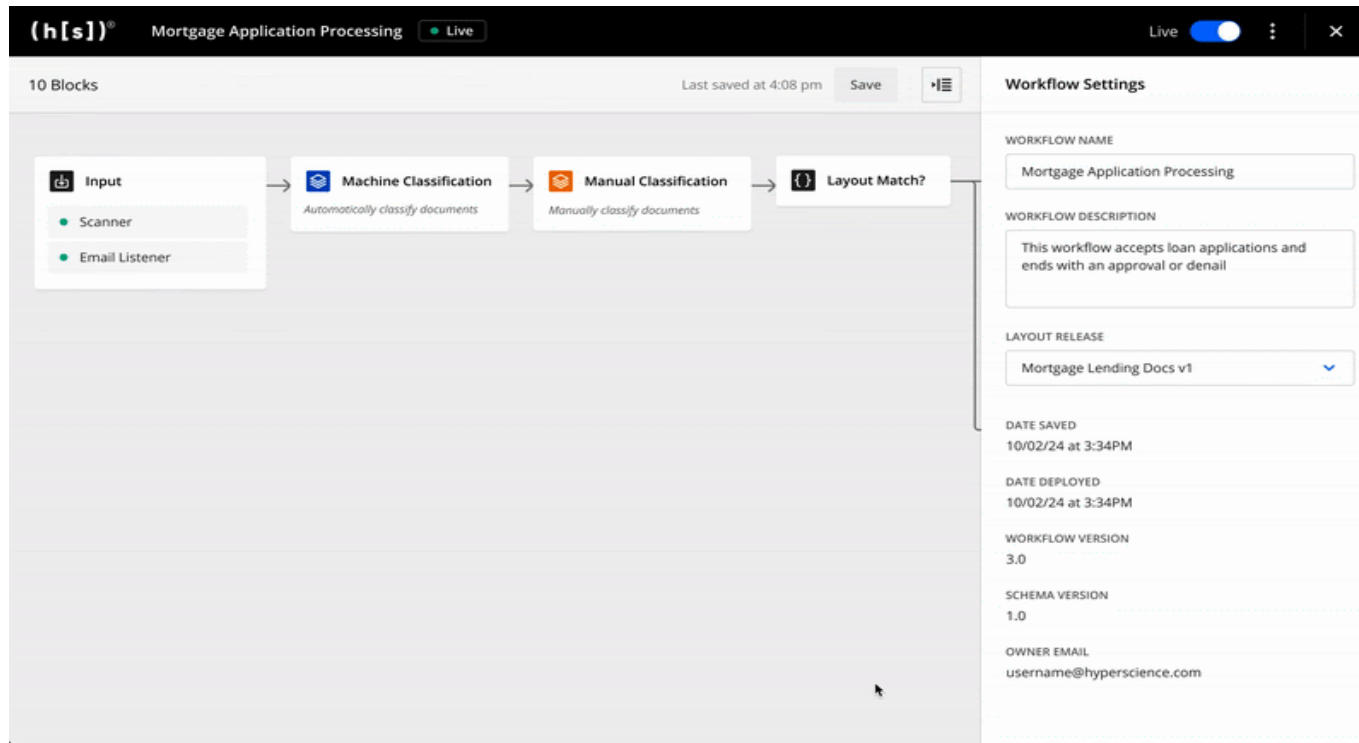
Best-In-Class Extraction Performance

내장된 QA 매커니즘은 시간이 지남에 따라 교정 정보를 반영하여 정확성을 높여주고 사람의 개입을 지속적으로 줄여 줄입니다.

이 "신뢰의 정확성"을 통해 고객은 사람의 개입없이 더 많은 정보를 추출 할 수 있습니다

차별화 기술 - 워크플로우

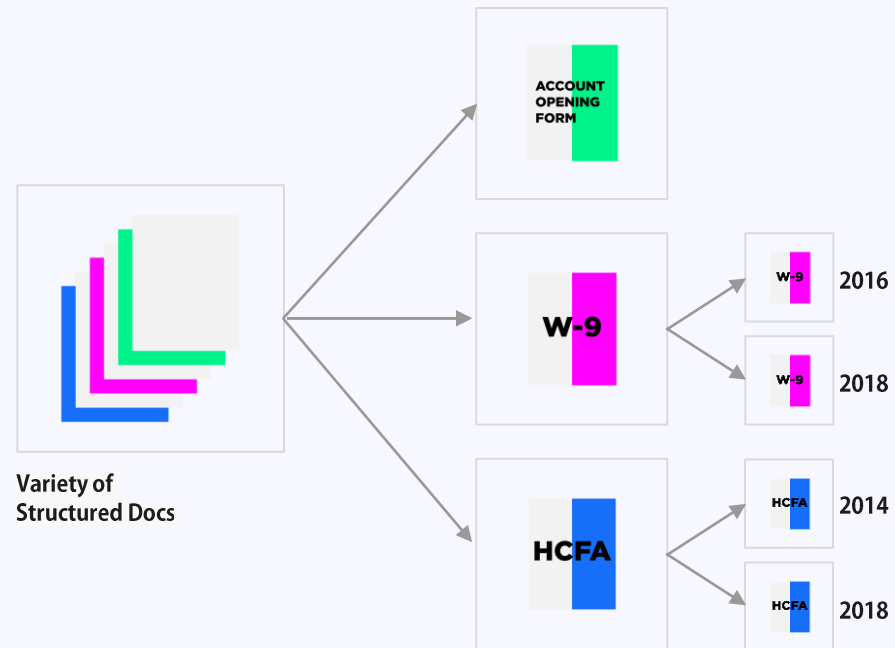
Skill 1



차별화 기술 - 자동분류

Skill 2

- 전처리(분류) 작업이 필요 없는 자동 문서 분류
- 문서들의 유사도 별로 그룹핑하여 맵핑

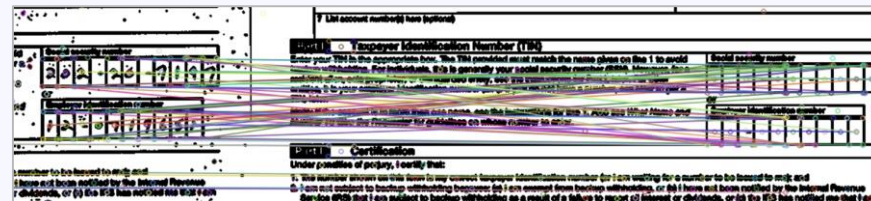
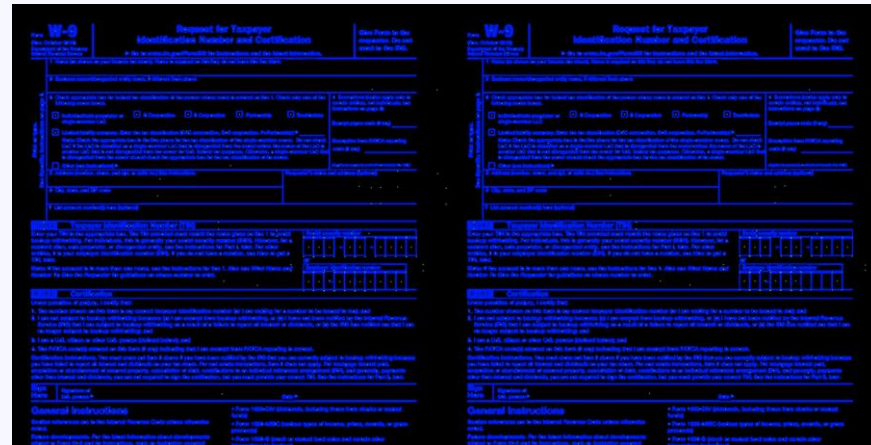


차별화 기술 – VPC(VISUAL PAGE CLASSIFIER)

Skill 3

- 문서를 올바르게 분류하는 것은 올바른 추출을 위해 가장 중요합니다.
- 양식에서 각 개체의 중간점을 찾아 이를 활용하여 왜곡 또는 늘어나거나 선명하지 않은 이미지의 문서라도 정확하게 문서를 분류해 내는 기술을 사용합니다.

Bottom line
문서들을 학습하여 빈 문서의 템플릿을 생성하여 문서를 자동 분류합니다.



차별화 기술 - 필드 속성 학습 AI

Skill 4

필요한 정보만 정확한 추출

불필요한 데이터는 제외하고
필요한 데이터만 정확하게 추출

문장 구조 AI 학습

주소 필드의 경우 구조를 학습해서
앞쪽에는 숫자가 나오고 중간에는
숫자가 나오지 않으므로 숫자 1과 매우 유사한
소문자 ll을 문자로 인식

필드의 속성 학습

글자를 인식하는 방식이 아닌
아닌 필드의 속성 및 의도를
학습하는 5년간 검증된 학습 엔진

박스를 벗어난 문자 인식

2 Business name/disregarded entity name, if different from above

General Electric

GENERAL ELECTRIC

5 Address (number, street, and apt. or suite no.)

123 Electron Still Hill St.

123 ELECTRON STILL HILL ST

2 Business name/disregarded entity name, if different from above

RWE ~~LLC~~ LLC

RWE LLC

Social security number

123-45-6789

123-45-6789

차별화 기술 - 해상도 비종속

Skill 5

Page 3 of 5

Form Approved
OMB No. 0900-0023

Form No. 1004-108 (Rev. 12-2010)

WHOSE RECORDS TO BE DISCLOSED
NAME: Johny Rocket
SSN: 123-46-7890
DATE: June 12, 1997

AUTHORIZATION TO DISCLOSE INFORMATION TO THE SOCIAL SECURITY ADMINISTRATION (SSA)
PLEASE READ THE ENTIRE FORM, BOTH PAGES, BEFORE SIGNING BELOW

I voluntarily authorize release of disclosure (including paper, oral, and electronic interchange) OF WHAT All my medical records; also education records and other information related to my ability to perform tasks. This includes specific permission to release:

- All records and other information regarding my treatment, hospitalization, and outpatient care for my impairment(s) including:
 - Psychological, psychiatric or other mental impairment(s) (including "psychotherapy notes" as defined in 45 CFR 164.501)
 - Drug abuse, alcoholism, or other substance abuse
 - Stable self reports
 - Records which may indicate the presence of a communicable or noncommunicable disease, and tests for records of HIV/AIDS
 - Gene-related impairments (including genetic test results)
- Information about how my impairment(s) affects my ability to complete tasks and activities of daily living, and affects my ability to work.
- Copies of educational tests or evaluations, including individualized Educational Programs, interest assessments, psychological and speech evaluations, and any other records that can help evaluate function; also teachers' observations and evaluations.
- Information created within 12 months after the date this authorization is signed, as well as asset information.

FROM WHOM

- All medical sources (hospitals, clinics, labs, physicians, psychologists, etc.) including mental health, corrective, addiction treatment, and VA health care facilities.
- All educational sources (schools, teachers, records administrators, etc.)
- Social workers/rehabilitation counselors
- Consulting examiners used by SSA
- Employers, insurance companies, workers' compensation programs
- Others who may know about my condition (family, neighbors, friends, public officials)

TO WHOM
The Social Security Administration and to the State agency authorized to process my case (usually called "disability determination services"), including contract copy services, and doctors or other professionals consulted during the process. Also, for informational details, to the U.S. Department of State Foreign Service Post.

PURPOSE
Determining my eligibility for benefits, including benefits at the discretion of any treatment(s) that I may receive and meet SSA's definition of disability; and whether I can manage such benefits.

EXPIRES WHEN
This authorization is good for 12 months from the date signed below my signature.

PLEASE SIGN USING BLUE OR BLACK INK ONLY (If not signed by subject of disclosure, specify basis for authority to sign)

SIGN [Signature]
Date: 6/12/2010
Street Address: 111 Illinois Avenue
City: Chicago
State: IL
ZIP: 60632

WITNESS
I know the person signing this form or am authorized to sign on their behalf.
SIGN [Signature]
Date: 6/12/2010
Street Address: 636-123-4567

2020-01-01 19:50:38 Danielle McKenzie Morris, Jones and Thoric

PHYSICIAN

GROUP NO. 231145 | SPEC. ED | PHYSICIAN OFFICE ADDRESS NO. 7-12317

PHYSICIAN NAME: Dr. Michael Gallaghy
EMAIL: drmg@yosemitephys.com
PHONE NO. 403-240-6648

DATE OF SERVICE: 17 3 20
PRECEDENCE CODE: D 2 3 9 8
MILE CODE: 1 2
MEDI: DMG
PHYSICIAN FEE: 915.00
LABORATORY CHARGE: 30.00
TOTAL CHARGE: 945.00

DATE OF SERVICE: 17 3 20
PRECEDENCE CODE: F 0 2 1 0
MILE CODE: 1 2
MEDI: N/A
PHYSICIAN FEE: 1250.00
LABORATORY CHARGE: 137.50
TOTAL CHARGE: 1387.50

DATE OF SERVICE: 24 3 20
PRECEDENCE CODE: 4 1 2 0 0
MILE CODE: 2 7
MEDI: N/A
PHYSICIAN FEE: 875.50
LABORATORY CHARGE: 25.00
TOTAL CHARGE: 900.50

ALLOWED AMOUNT	MC	%	PHYSICIAN SHARE
500.00	Y	53	445.00
1000.00	N	72	375.00

CHECK NO. 1214561 DATE: 3/4/20
PRODUCED: 750.00 PHYSICIAN FEE: 679.00
CLAIM NO. 881 0024561190

TOTAL FEE SUBMITTED: 3,233.00

INSTRUCTIONS FOR CLAIM SUBMISSION
GROUP NO. 1223371 SUBSCRIPTION NO. 896
EMPLOYER: Stamper Construction
NAME OF INSURING AGENT OR PLAN: Yosemite Insurance

YOUR NAME (PLEASE PRINT): Frederic J. LeClair
YOUR CERT. NO. OR L.A. NO.: 213 920 891
YOUR DATE OF BIRTH: 10 4 1978

RELATIONSHIP TO EMPLOYER/PLAN: Parent
DATE OF BIRTH: 5 2 03
IF STUDENT, INDICATE SCHOOL: Univ. of Ottawa
PRESENT L.A. NO.: 121 3278118

DATE OF SERVICE: 6 4 20
PHYSICIAN FEE: 239482
NAME OF OTHER INSURING AGENT OR PLAN: [Signature]

DATE OF SERVICE: 01 03 20
DATE OF SERVICE: 05 03 20
A CONTRACT RALEN: 01 01 20
NAME OF EMPLOYEE: Henry Jones
POSITION TITLE: claims representative

차별화 기술 - 뛰어난 수기 인식

Skill 6

◀ | ← 1 → of 3 | 🔍 100% 🔍 | 🖨 ⌵

은행거래신청서 신한은행 Shinhan Bank

본인은 신한은행의 예금(신탁)거래기본약관, 취급(신)별 약관 및 해당거래약관에 적용됨을 승인하고 아래의 같이 거래를 신청합니다.
 * '특정금융거래정보보고및이용등이관한법률'에 따라 고객이 금융기관에 거래시에는 아래의 내용을 기재하여야 하며 제공하신 정보는 특별에 의거 중도 외의 목적으로 사용되지 않습니다.
 금융사주동사명 제8호의 국제 근거지에 신청내용 그룹내 자회사명에는 고객관계적 제공 및 이용이 가능하며, 고객정보 취급방침에 적용 설명 및 조부를 확인합니다.

이수연 1982년 05월 07일
 대한민국
 서울특별시 서초구 신반포로 33길 67
 서울특별시 영등포구 여의도동 82길 36
 02 3567 1549 인감
 070 2676 4302
 010 5741 3219 서명

1-가입상품명
 Normalized
 Raw
 Identified by
 Transcribed by

자택	● 서울시 서초구신반포로 33길 67	>
직장	● 서울시 영등포구여의도동 82길 36	>
휴대폰	010 5741 3219	>
서명	Signed	>
급여소득자	Not Checked	>
전문직	Checked	>
자영업자	Not Checked	>
연금소득자	Not Checked	>
직장명	하이퍼사이언스	>
부서명	프로젝트팀	>
1-가입상품명	<ul style="list-style-type: none"> Normalized Raw Identified by Transcribed by 	<ul style="list-style-type: none"> 숄정기적금 숄정기적금 Machine Machine

차별화 기술 - 뛰어난 수기 인식(계속)

Skill 7

3 of 8

100%

PAGE 3

모두가 올바른 정보를 주어 이해하였으며, 농업 제3종 제6항에 따라 설명 들었음을 확인합니다.

위의 내용을 설명 들었음

원화 입출금예금 작성 항목

통장, 현금카드를 타인에게 압도하는 경우 손해배상책임을 부담할 수 있고, 전자금융거래법에 의해 처벌받을 수 있습니다. 또한 입출금이 자유로운 예금약관에 따라 계좌개설 등의 금융거래가 제한될 수 있음에 대해 설명 들었음

입출금예금(외화예금 포함) 추가 작성 항목

거래목적	<input type="checkbox"/> 저축/투자 <input type="checkbox"/> 급여 및 생활비 <input type="checkbox"/> 결제계좌(카드·광과금·대출금·보험료) <input type="checkbox"/> 자금이체 <input type="checkbox"/> 부채/이차상환 <input type="checkbox"/> 상속/증여성거래 <input type="checkbox"/> 여행경비 <input type="checkbox"/> 유학경비 <input type="checkbox"/> 해외체재비 <input type="checkbox"/> 해외이주비 <input type="checkbox"/> 기타(기타: 구체적으로 기재할수) ※ 개인사업자 또는 법인의 경우 아래 항목 중 해당하는 거래목적 추가 작성(선택지제) <input type="checkbox"/> 사업상거래(물품대금결제) <input type="checkbox"/> 자금관리 <input type="checkbox"/> 수수료 <input type="checkbox"/> 해외직접투자 <input type="checkbox"/> 무역거래 <input type="checkbox"/> 무역외거래
월별 예상거래 금액	<input type="checkbox"/> 1백만원 이하 <input type="checkbox"/> 5백만원 이하 <input type="checkbox"/> 2천만원 이하 <input type="checkbox"/> 1억원 이하 <input type="checkbox"/> 1억원 초과
월별 예상거래 건수	<input type="checkbox"/> 10회 이하 <input type="checkbox"/> 50회 이하 <input type="checkbox"/> 100회 이하 <input type="checkbox"/> 100회 초과

정기예금(외화 포함), 정기적금 작성 항목

1. 은행상품 구속행위 규제제도 관련 유의사항 안내
 본인이 중소기업의 대표자이거나 신용평점이 낮은 개인(개인신용평점 기준 하위 100분의 10)인 경우, 은행법(제52조)의상 구속행위 여부 판정에 따라 아래 거래가 제한될 수 있습니다.
 ① 수신상품의 신규
 ② 수신상품 신규일 이후 1개월 이내 본인 명의 대출 거래(중소기업, 개인신용평점 기준 하위 100분의 10에 해당하는 개인)
 ③ 본인이 재직 중인 중소기업의 대출 거래(중소기업의 대표자)
 상기 내용에 대해 설명 들었음

2. 정기예금 제예치 신청서
 ① 대상예금: 퍼스트정기예금(종별 96 중 12개월제, 단, 비거주자/비과세종합저축/법인/임의단체는 제외)
 ② 만기제예치 요건
 - 최장 5년 이내에서 매 만기일에 원금에 세후이자를 합산하여 동일기간으로 만기일 연장
 ③ 적용금리: 만기일(재신규일) 시점에 영입점 및 인터넷홈페이지에 게시된 해당 종별 및 기간별 고시이율
 상기 내용에 대해 설명 들었음

기타 작성 항목

■ 예금잔액조회서 특보
 작성 대상 : 입출금예금, 정기예금 5만원 이상, 정기적금(자유정리식)

예금보험관계-위의 내용	Checked
예금보험관계 설명 확인	● 설명 들었음
금융실명거래 확인 작성 확인-위의 내용	Checked
금융실명거래 확인-설명들었음	● 설명 들었음
원화입출금예금 작성항목 확인-설명들었음	● Supervision Required, But Disabled Supervision was required for this field, but supervision was disabled for the submission.
Normalized	설명 들었음
Raw	설명 들었음
Identified by	Machine
Transcribed by	Machine

차별화 기술 - QA를 통한 정확도 지속적인 상승

Skill 8

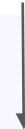
- 내재된 샘플링 검사 및 QA를 통한 학습으로 정확도 향상
- 필드 레벨의 Confidence 제공
- MVT (Model Validation Task)
- 기계와 휴먼의 최적 조합을 극대화한
Human In The Loop

HINESTANNER@WHEELER.COM



HINESTANNER@WHEELER.COM

Teacher special educational needs



Teacher special educational needs

주요 차이점-IDP Vs. OCR Vs. ICR

Categories	Hyperscience IDP	OCR	ICR
Text Types	Handwritten text, cursive, machine-typed	Machine-typed text	Handwritten text
Text Parameters	Free form, cursive	Constrained - characters separated into individual boxes	
Extraction	Field level with context	Character by character with no context	
Accuracy Measurement	Field level	Character level	
Image Quality	< 100 DPI	300 DPI +	
Classification	Machine Learning tackles low DPI, distorted/skewed images	Rules-Based - struggles with low DPI or other common doc imperfections	
Image Correction	Yes	No	
Pre-Sorting	No	Yes - needs separate engines for print + handwriting	

비교

	Hyperscience	Legacy Tech	Big Tech	RPA OCR
	(h[s]) [®]	ABBYY opentext IBM KOFAX [®]	G Microsoft Azure aws	AUTOMATION ANYWHERE UiPath blueprism
Classification	✓	—	✗	—
Extraction	✓	—	—	—
Built-In Quality Assurance	✓	—	✗	—
Target Accuracy	✓	✗	✗	✗
Human-in-the-Loop	✓	—	✗	—
Reporting	✓	—	✗	—

하이퍼사이언스는 지능형 문서인식 (IDP: Intelligent Document Processing) 시장의 새로운 강자입니다.



BOND

FIRSTMARK

TIGERGLOBAL



McKinsey & Company

\$300M

3억불 투자 유치

300+

글로벌 임직원

뉴욕, 런던, 소피아, 베를린,
민헨, 토론토, 호주, 인도,
한국



accenture

WNS



aws

Microsoft



Fidelity

AIG



Guardian

QBE

Everest Group
(21)

IDP PEAK Matrix
Major Contender

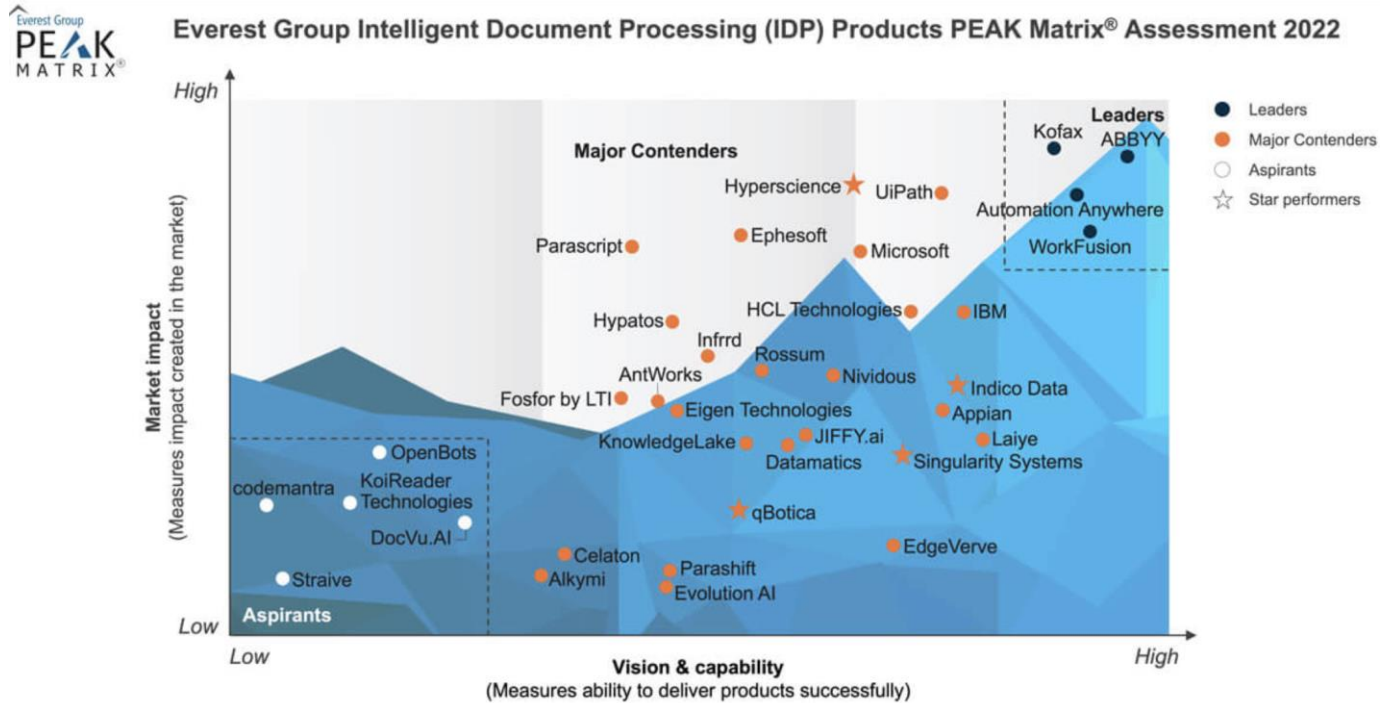
CB
Insights (21)

AI 100

*Preliminary, confidential draft for discussion

IDP 2022

PEAK Matrix

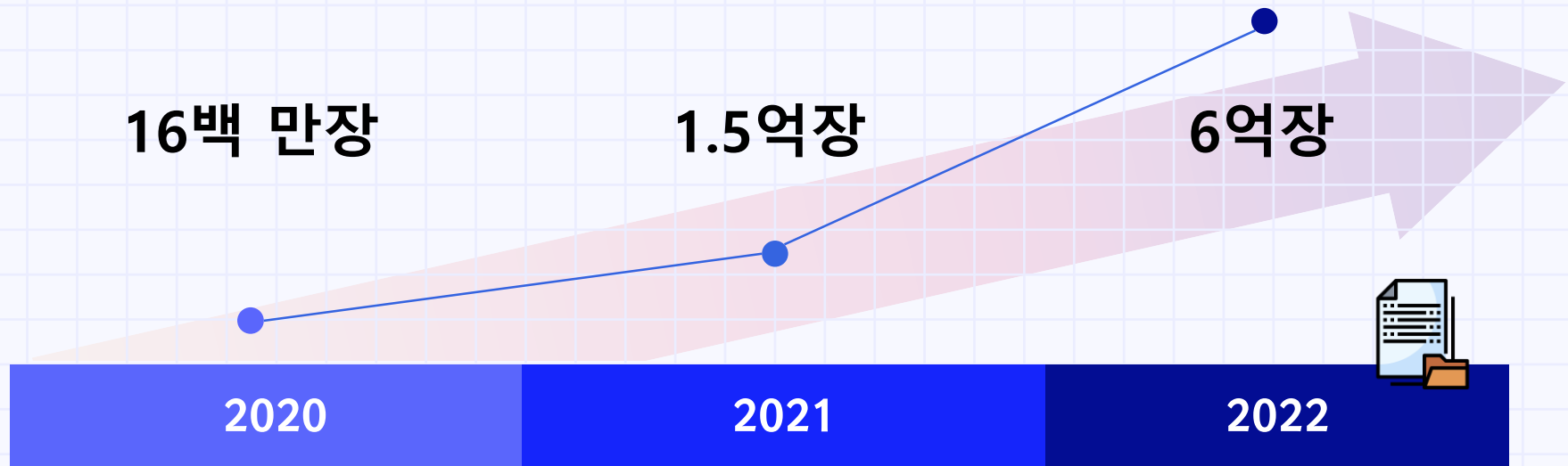


Source: Everest Group (2022)

고객 성공 사례



고객 문서 처리 건수




2년 37.5배 증가



Thanks for **Your** **Attention**

Always will do our best!

 누리어시스템