

AWS 기반 보이스피싱 탐지·분석 시스템



팀명: 피스메이커

팀원: 정상호, 이의석

멘토: 홍대화(SK 실더스), 유일해(원티드랩)

목차

1. 팀 소개 & 일정
2. 주제 선정 배경
3. 프로젝트 목표
4. 시스템 설계
5. 프로젝트 결과
6. 소회
7. Q&A

팀 소개 & 일정

팀명

팀소개

타임라인

팀명

팀 소개 & 일정
팀명

PEACEMAKER

보이스피싱 없는
안전한 통화 환경을 만들자

팀 소개 & 일정

팀 소개 & 일정
팀소개

정상호

데이터 수집·변환 담당

- Amazon conntect로 통화 자동 녹음
- 녹취 파일 S3(Recordings)에 저장
- Lambda A를 통하여 Transcribe Batch Job 실행
- 텍스트 변환 결과를 관리

이의석

데이터 분석·알림 담당

- Transcribe 결과를 S3(Results)에 저장
- Lambda B에서 Comprehend 호출해 키워드·감정·엔터티 분석
- 위험도 임계치 초과 시 SNS로 알림 발송

공동

데이터 시각화 공동 담당

- Athena 테이블 설계 및 쿼리 작성
- QuickSight 대시보드 제작
- 위험 키워드 발생 빈도 통화별 위험도 분포 시각화

타임라인

팀 소개 & 일정
타임라인

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31	1					

08.18 ~ 08.20
주제선정 & 자료수집

08.21 ~ 08.27
기능 구현 & 테스트

08.26 ~ 09.01
PPT & 발표 준비

주제 선정 배경

피해 사례

선정 이유

피해 사례

주제 선정 배경
피해 사례

B씨는 저금리 대출로 갈아
이라는 사람이 보내준 어플
거래 은행에서 어제 신청한
B씨는 급히 돈을 빌려 대면
대면편취로 금감원에 환급

2025년 1분기 보이스피싱 피해액 3116억원...2024년 대비 2.2배 급증

세계일보 원문 | 기사전송 2025-04-27 15:21

댓글 0

보도 0

AI챗으로 요약

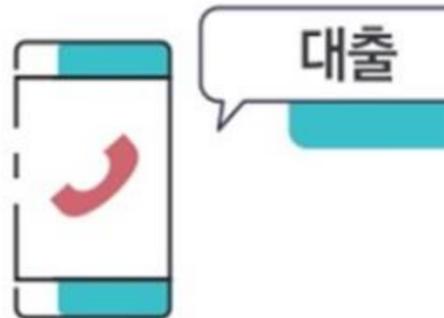


50대 이상 피해자 비중 53% 차지

악성앱 통한 개인정보 탈취 주의해야

올해 1~3월 보이스피싱(전화금융사기) 범죄가 5878건 발생해 피해액이 3116억원에 달한 것으로 나타났다. 전년 동기 대비 범죄 건수는 17%, 피해액은 2.2배 늘었다. 건당 평균 피해액 역시 5301만원으로 지난해(2813만원) 대비 1.9배 증가했다.

사칭형' 범죄가 전체
%에서 올해 10%포



주제 선정 이유

주제 선정 배경
선정 이유

금융사기 사건 급증으로 통화 데이터 분석 기반 보안 솔루션 중요성 대두

고도화된 범죄 수법 확산으로 자동화된 위협 탐지 시스템 개발 필요성 증가

실시간 위협 탐지와 자동 대응으로 능동적 보안 시스템 도입 필요성 증대

프로젝트 목표

목표

기대효과

프로젝트 목표

프로젝트 목표
목표

보이스피싱 대응을 위한 분석 기반 음성 데이터 보안 파이프라인 구현

분석 결과 데이터베이스화로 추세 분석 및 시각화 지원

음성 → 텍스트 변환으로 검색·분석 가능한 형태 가공

자연어 분석을 통한 위험 키워드와 감정 패턴 탐지

기대효과

프로젝트 목표
기대효과

실시간 위험 키워드 탐지로 보이스피싱 피해 사전 예방 및 차단

자동 분석을 통한 연간 수천억원 규모 금융사기 피해액 감소 기여

데이터 기반 위험 패턴 분석으로 범죄 대응력 강화 및 재발 방지

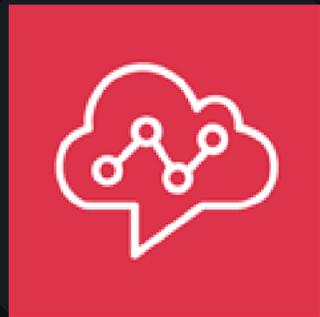
시스템 설계

기술 스택

아키텍처

기술 스택

시스템 설계
기술 스택



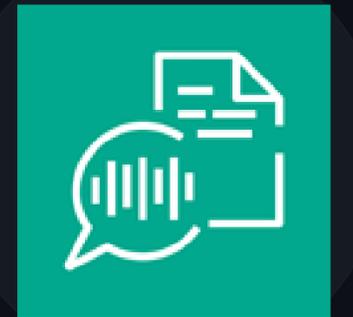
CONNECT



S3



LAMBDA



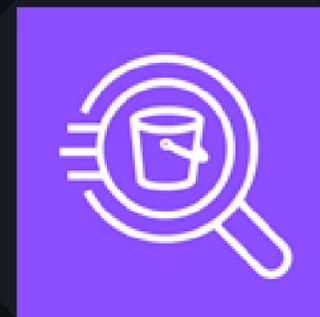
TRANSCRIBE



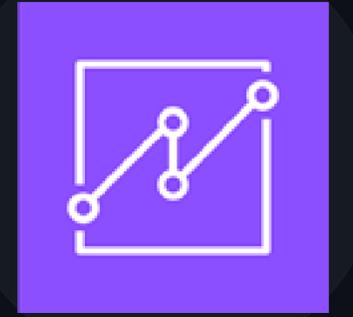
COMPREHEND



SNS



ATHENA



QUICKSIGHT



CONNECT

- Contact Flow로 통화 흐름 자동 제어
- 통화 종료와 동시에 녹음 파일 S3 자동 전송



S3

- 통화 녹음과 분석 결과를 안전하게 분리 저장, 버킷별 보안 강화
- 새 파일 업로드 감지로 후속 Lambda 작업 자동 시작



LAMBDA

- 파일 업로드 감지해 즉시 처리 시작, 서버리스 확장 가능한 분석
- 비용 최적화된 이벤트 기반 실행



TRANSCRIBE

- 통화 참여자 구분 및 위험 발언 시간 기록
- 배치 처리로 높은 정확도의 텍스트 변환



COMPREHEND

- "계좌이체", "보증금" 등 위험 단어 자동 탐지 및 감정 분석
- 전화번호, 계좌번호 등 민감정보 자동 마스킹



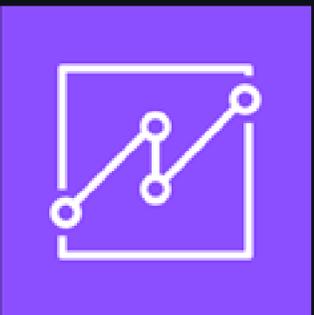
SNS

- 위험 통화 발견 즉시 관리자에게 다중 채널 알림
- 위험도에 따른 알림 우선순위 차등 적용



ATHENA

- 수천 건의 통화 데이터를 SQL로 효율적 검색 및 분석
- 월별, 시간대별 보이스피싱 발생 패턴 추출

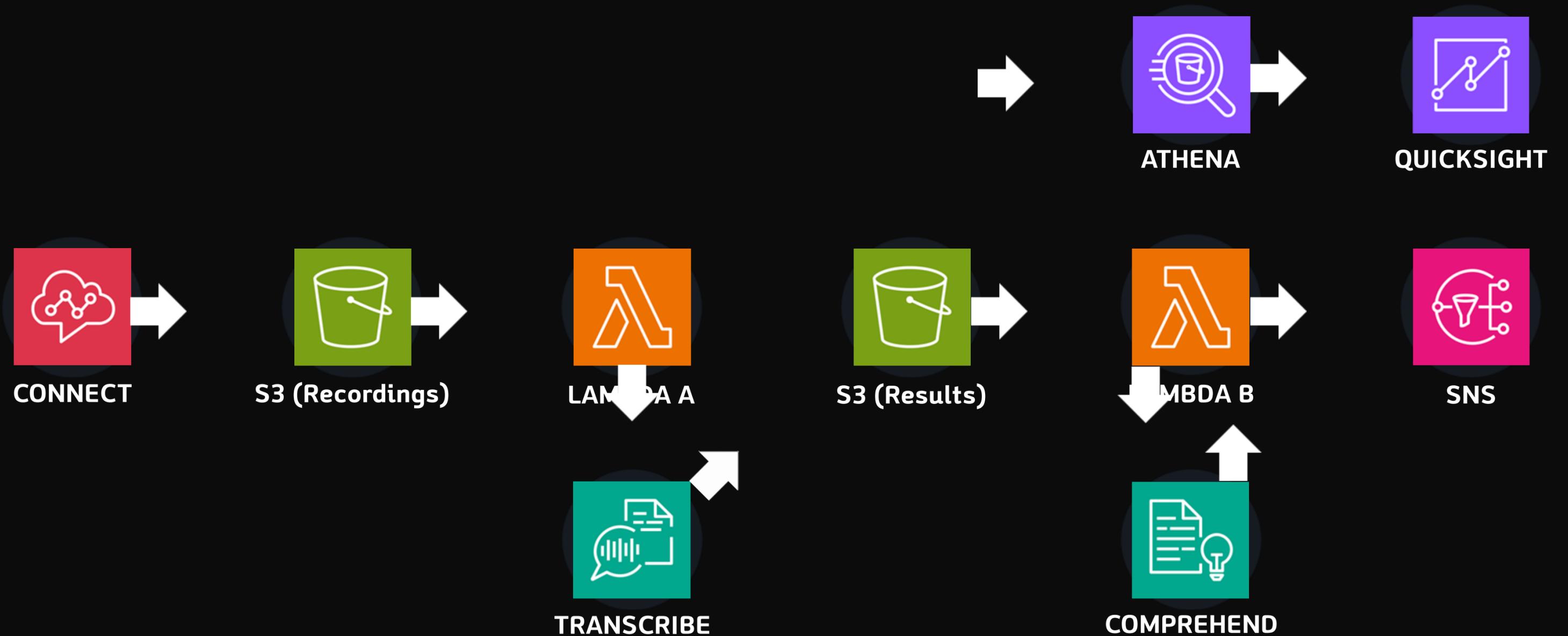


QUICKSIGHT

- 관리자용 실시간 모니터링 대시보드 구성
- 실시간 데이터 업데이트 자동 반영으로 차트 시각화

아키텍처

시스템 설계
아키텍처



프로젝트 결과

구현

시연 영상

구현

프로젝트 결과 구현

항목

recordings-d0cef234 > raw/ > / > ivr/ > 2025/ > 08/ > 31/

31/ S3 URI 복사

객체 속성

객체 (3) 🔄 S3 URI 복사 URL 복사 다운로드 열기 삭제 작업 폴더 만들기 업로드

객체는 Amazon S3에 저장되어 있는 기본 엔터티입니다. [Amazon S3 인벤토리](#)를 사용하여 버킷에 있는 모든 객체의 목록을 얻을 수 있습니다. 다른 사용자가 객체에 액세스할 수 있게 하려면 명시적으로 권한을 부여해야 합니다. [자세히 알아보기](#)

🔍 접두사로 객체 찾기

<input type="checkbox"/>	이름	▲	유형	▼	마지막 수정	▼	크기	▼	스토리지 클래스	▼
<input type="checkbox"/>	679a63b4-5891-4133-ac2f-c9c89dc51b05_20250831T09:07_UTC.wav		wav		2025. 8. 31. pm 6:08:07 PM KST		1.5MB		Standard	

성공

오류

성공

오류

연결 해제

종료 이벤트

녹음 및 분석 동작 설정

에이전트 및 고객 음성 녹음...

전화번호로 전송

출력: +821037743729

구현

프로젝트 결과 구현

```
def analyze_with_comprehend(text: str, lang: str) -> Dict[str, Any]:
    chunks = _chunk_text(text)
    sentiments, key_phrases, entities, pii_entities = [], [], [], []
    lc = lang if lang != "auto" else "en"
    for ch in chunks:
        r, _ = _try(comprehend.detect_sentiment, Text=ch, LanguageCode=lc)
        if r: sentiments.append(r)
        r, _ = _try(comprehend.detect_key_phrases, Text=ch, LanguageCode=lc)
        if r: key_phrases.extend(r.get("KeyPhrases", []))
        r, _ = _try(comprehend.detect_entities, Text=ch, LanguageCode=lc)
        if r: entities.extend(r.get("Entities", []))
        r, _ = _try(comprehend.detect_pii_entities, Text=ch, LanguageCode=lc)
        if r: pii_entities.extend(r.get("Entities", []))
    overall = None
    if sentiments:
        labels = [s.get("Sentiment") for s in sentiments if s.get("Sentiment")]
        if labels: overall = max(set(labels), key=labels.count)
    return {"language": lang, "overall_sentiment": overall,
            "key_phrases": key_phrases, "entities": entities, "pii_entities": pii_entities}
```

```
def score_risk(text: str, analysis: Dict[str, Any]) -> Tuple[float, List[str]]:
    lower, hits = text.lower(), []
    for kw in DANGEROUS_KEYWORDS:
        if kw.lower() in lower: hits.append(kw)
    pii_count = len(analysis.get("pii_entities", []))
    neg = 1.0 if (analysis.get("overall_sentiment") or "").upper() == "NEGATIVE" else 0.0
    score = KEYWORD_WEIGHT * len(hits) + PII_WEIGHT * pii_count + NEGATIVE_SENTIMENT_WEIGHT * neg
    return score, {"keyword_hits": hits, "pii_count": pii_count, "negative_sentiment": bool(neg)}
```

transcribe-results-d0cef234 > analytics/

analytics/

S3 URI 복사

객체 속성

객체 (2)

[S3 URI 복사](#) [URL 복사](#) [다운로드](#) [열기](#) [삭제](#) [작업](#) [폴더 만들기](#) [업로드](#)

객체는 Amazon S3에 저장되어 있는 기본 엔터티입니다. [Amazon S3 인벤토리](#)를 사용하여 버킷에 있는 모든 객체의 목록을 얻을 수 있습니다. 다른 사용자가 객체에 액세스할 수 있게 하려면 명시적으로 권한을 부여해야 합니다. [자세히 알아보기](#)

접두사로 객체 찾기

<input type="checkbox"/>	이름	▲	유형	▼	마지막 수정	▼	크기	▼	스토리지 클래스	▼
<input type="checkbox"/>	connect-679a63b4-5891-4133-ac2f-c9c89dc51b05_20250831T09-07_UTC.wav-1756631287-f91e4d.summary.json		json		2025. 8. 31. pm 6:08:16 PM KST		2.2KB		Standard	

구현

프로젝트 결과 구현

```
for kw, w in KW_WEIGHTS.items():
    pat = patterns.get(kw, re.escape(kw))
    flags = re.IGNORECASE if kw in ("OTP", "URL", "CVC") else 0
    try:
        matches = re.findall(pat, text, flags)
    except re.error:
        matches = re.findall(re.escape(kw), text)
    cnt = len(matches)

    if cnt > 0:
        hits_unique.append(kw)
        hit_counts[kw] = cnt
        if KEYWORD_COUNT_DUPLICATES:
            mult = min(cnt, KEYWORD_MAX_PER_KEY) # per-key 상한 적용
        else:
            mult = 1
        numerator += (float(w) if USE_EXPLICIT_WEIGHTS else float(KEYWORD_WEIGHT)) * mult

denom = (sum(float(v) for v in KW_WEIGHTS.values()) if USE_EXPLICIT_WEIGHTS
         else max(1.0, len(KW_WEIGHTS) * float(KEYWORD_WEIGHT)))

if KEYWORD_CAP_AT_MAX:
    numerator = min(numerator, denom)

pct = max(0.0, min(100.0, (numerator / denom) * 100.0))

return {
    "score_raw": round(numerator, 2),
    "score_max": round(denom, 2),
    "score_pct": round(pct, 1),
    "keyword_hits": hits_unique,
    "keyword_counts": hit_counts,
    "keyword_hits_verbos": [f"{k} x{hit_counts[k]}" for k in hits_unique]
}
```

```
"risk_score_pct": 100.0,
"risk_detail": {
    "score_raw": 100.0,
    "score_max": 100.0,
    "score_pct": 100.0,
    "keyword_hits": [
        "검찰",
        "경찰",
        "세금",
        "압류",
        "인증번호",
        "보안카드",
        "비밀번호",
        "계좌",
        "원격",
        "앱설치",
        "링크",
        "카드번호"
    ],
    "keyword_counts": {
        "검찰": 3,
        "경찰": 2,
        "세금": 2,
        "압류": 1,
        "인증번호": 2,
        "보안카드": 1,
        "비밀번호": 1,
        "계좌": 1,
        "원격": 1,
        "앱설치": 1,
        "링크": 1,
        "카드번호": 1
    },
    "pii_count": 0,
    "negative_sentiment": false
},
```

```
"analysis": {
    "language": "ko",
    "overall_sentiment": "NEUTRAL",
    "key_phrases": [
        {
            "Score": 0.9969513416290283,
            "Text": "검찰",
            "BeginOffset": 0,
            "EndOffset": 2
        },
        {
            "Score": 0.975699782371521,
            "Text": "경찰",
            "BeginOffset": 4,
            "EndOffset": 6
        },
        {
            "Score": 0.9952965378761292,
            "Text": "세금",
            "BeginOffset": 8,
            "EndOffset": 10
        },
        {
            "Score": 0.9944043159484863,
            "Text": "헤나",
            "BeginOffset": 12,
            "EndOffset": 14
        },
        {
            "Score": 0.96739661693573,
            "Text": "압류, 인증번호 보안카드",
            "BeginOffset": 16,
            "EndOffset": 29
        },
        {
            "Score": 0.813785195350647,
            "Text": "비밀번호 검찰 계좌, 전화번호 원격",
            "BeginOffset": 40,
            "EndOffset": 59
        }
    ]
},
```

구현

프로젝트 결과 구현

```
caller_raw = numbers.get("caller")
dialled_raw = numbers.get("dialled")
caller_for_subject = _mask_e164(caller_raw) if (SUBJECT_MASK_NUMBERS and caller_raw) else (caller_raw or "")
dialled_for_subject = _mask_e164(dialled_raw) if (SUBJECT_MASK_NUMBERS and dialled_raw) else (dialled_raw or "")
```

```
phrases = top_items(ana
if phrases:
    lines.append("주요
for p in phrases:
```

[Connect] 위험 100% • ko • 발신 +8210*162** ▶ 받은편지함 x

yonsai 나에게 ▾ 오후 6:08 (26분 전) ☆ 😊

📞 Amazon Connect 통화 분석 결과

- 착신번호: +12132910431
- 발신번호: +821048196162

- 위험도: ● HIGH (100/100 = 100.0%)
- 언어: 🇰🇷 ko
- 적중 키워드(12): 검찰 x3, 경찰 x2, 세금 x2, 압류 x1, 인증번호 x2, 보안카드 x1, 비밀번호 x1, 계좌 x1, 원격 x1, 앱설치 x1, 링크 x1, 카드번호 x1
- 주요 키워드:
 - 검찰
 - 세금
 - 해나
 - 경찰
 - 압류, 인증번호 보안카드

- 권장 조치:
- 고위험 신호입니다.
 - 결제/인증·개인정보 교환을 즉시 중단하세요.
 - 보안 담당자에게 즉시 공유하세요.

요약 스니펫:
검찰, 경찰, 세금, 해나, 압류, 인증번호 보안카드, O. T. P. 비밀번호 검찰, 계좌, 전화번호 원격? 카드 번호 앱 설- 앱 설치, 링크 인증번호 경찰 세금 검찰

<  **00350447873077777** ▾ ⋮

저장되지 않은 번호에서 온 메시지입니다. 스미싱이나 피싱에 유의하세요.

수신 차단

메시지 신고

8월 31일 일요일

[국외발신]

● 보이스피싱 의심 100% — 통화 즉시 중단.
<https://www.counterscam112.go.kr>

오후 6:08

구현

프로젝트 결과 구현

QuickSight

< 데이터 세트

Connect Summary Dataset

데이터

요약 새로 고침 권한 사용량

소개

SPICE 크기: 3.6KB

새로 고침

✓ 상태 완료함 6개 행 가져오기(100% 성공)

🕒 마지막으로 성공한 새로 고침 2025년 9월 1일 오후 2:00 GMT+9

액세스 설정

공유: 소유자(1) 최종 사용자(0)

행 수준 보안: = 제한 없음 [설정](#)

열 수준 보안: || 제한 없음 [설정](#)

스키마

고유 키 [자세히 알아보기](#)

비활성화됨

데이터 세트

이름	소유자	최근 수정
Connect Summary Dataset	나	a few seconds ago

새 데이터 세트

구현

프로젝트 결과 구현

QuickSight
🏠 ?

분석 및 기타 정보 찾기 🔍

- ★ 즐겨찾기
- 🕒 최근
- 📁 내 폴더
- 👥 공유 폴더
- 📊 대시보드
- 📖 Data stories
- 📊 분석
- 📁 데이터 세트
- 🗣 커뮤니티
- 🏷 주제

대시보드

📊 대시보드

Connect Summary

업데이트: 3 days ago

일자별 평균 위험도(%)

risk_score_pct (평균)

day

위험도 밴드 분포

Risk Band

- HIGH
- MEDIUM
- LOW

그룹화 기준: risk_band
크기: risk_score_pct (개수)

언어별 건수

language

ko

risk_score_pct (개수)

최신 고위험 상세

created_at_ts	risk_band	language	hits_str	source_object
9월 1, 2025 12:00오전	HIGH	ko	경찰, 경찰, 인종번호, OTP, 보안카드, 비밀번호, 계좌, 계좌정보, 이체, 송금, 원격, 연결지, 링크	s3://transcribo-resul
9월 1, 2025 12:00오전	HIGH	ko	경찰, 인종번호, 보안카드, 비밀번호, 계좌, 계좌정보, 이체, 송금, 원격, 연결지, 링크	s3://transcribo-resul
9월 1, 2025 12:00오전	MEDIUM	ko	경찰, 보안카드, 비밀번호, 계좌, 계좌정보, 송금, 원격, 연결지, 링크	s3://transcribo-resul
9월 1, 2025 12:00오전	MEDIUM	ko	경찰, 비밀번호, 계좌, 계좌정보, 송금, 원격, 연결지, 링크	s3://transcribo-resul
8월 31, 2025 12:00오전	HIGH	ko	경찰, 경찰, 세금, 압류, 인종번호, 보안카드, 비밀번호, 계좌, 원격, 연결지, 링크, 카드번호	s3://transcribo-resul
8월 31, 2025 12:00오전	LOW	ko		s3://transcribo-resul

```

"ExportToCSVOption":
"SheetControlsOption"
    }
    )
    
```

시연 영상

공공기관 사칭 보이스피싱 시연

프로젝트 결과
시연 영상



보이스피싱



피해자

참고



경찰청마약조직범죄수사과



금융감독원



한국인터넷진흥원(KISA)

소회

대한민국(KR)

전화번호 주문

지원되는 리전	번호 유형	ID 요구 사항 유무	사용
아시아 태평양(서울)	VoIP 앞자리 번호: +82 70	예	현재 지원
	대표 번호: +82 15, +82 16	예	대표 지원
	수신자 부담 앞자리 번호: +82 80	예	한국 현재 지원
	지역 앞자리 번호: +82 2	예(포팅을 통해)	VO 그리션 Am 습

전화번호 이동

지원되는 리전	번호 유형	이동성 기간	필수 문서
아시아 태평양(서울)	지역 앞자리 번호: +82 2(+821, +825, +827, +82308 이외의 모든 +82 번호)	월요일~금요일 오전 9시 ~오후 6시 KST	<p>기존 번호에 대한 새 SIP 주문 양식 및 SIP 포팅 양식. 신청 시 제공된 양식을 사용하세요. 문서에 서명하는 직원의 생년월일이 기재되어 있어야 하며 회사 도장을 찍어야 합니다. 이 양식과 함께 다음과 같은 서류가 필요합니다.</p> <ol style="list-style-type: none"> 6개월 이내의 한국 기업 증명서. 증명서는 지역이 일치해야 합니다. 양식에 서명하는 사람의 개인 ID 인감 증명서 <p>지원 티켓을 제출하여 번호(들)의 이동성을 확인합니다.</p>
	국가 앞자리 번호: +82 50	월요일~금요일 오전 9시 ~오후 6시 KST	<p>기존 번호에 대한 새 SIP 주문 양식 및 SIP 포팅 양식. 신청 시 제공된 양식을 사용하세요. 문서에 서명하는 직원의 생년월일이 기재되어 있어야 하며 회사 도장을 찍어야 합니다. 이 양식과 함께 다음과 같은 서류가 필요합니다.</p> <ol style="list-style-type: none"> 6개월 이내의 한국 기업 증명서. 증명서는 지역이 일치해야 합니다. 양식에 서명하는 사람의 개인 ID 인감 증명서 <p>지원 티켓을 제출하여 번호(들)의 이동성을 확인합니다.</p>
	대표 번호: +82 15, +82 16	월요일~금요일 오전 9시 ~오후 6시 KST	<p>RN/TFN 변경 양식이 필요합니다. 신청 시 제공된 양식을 사용하세요. 이 양식과 함께 다음과 같은 서류가 필요합니다.</p> <ol style="list-style-type: none"> 6개월 이내의 회사 인증 등록증 사본 <p>지원 티켓을 제출하여 번호(들)의 이동성을 확인합니다.</p>
	수신자 부담 앞자리 번호: +82 80	월요일~금요일 오전 9시 ~오후 6시 KST	<p>RN/TFN 변경 양식이 필요합니다. 신청 시 제공된 양식을 사용하세요. 이 양식과 함께 다음과 같은 서류가 필요합니다.</p> <ol style="list-style-type: none"> 6개월 이내의 회사 인증 등록증 사본 <p>지원 티켓을 제출하여 새 번호(들) 주문에 대한 문서를 확인합니다.</p>
	VoIP 앞자리 번호: +82 70	월요일~금요일 오전 9시 ~오후 6시 KST	<p>사실상 다른 +8270으로의 통화 전달이 가능합니다.</p> <p>기존 번호에 대한 새 SIP 주문 양식 및 SIP 포팅 양식. 신청 시 제공된 양식을 사용하세요. 문서에 서명하는 직원의 생년월일이 기재되어 있어야 하며 회사 도장을 찍어야 합니다. 이 양식과 함께 다음과 같은 서류가 필요합니다.</p> <ol style="list-style-type: none"> 6개월 이내의 한국 기업 증명서. 증명서는 지역이 일치해야 합니다. 양식에 서명하는 사람의 개인 ID. 인감 증명서 <p>지원 티켓을 제출하여 번호(들)의 이동성을 확인합니다.</p>



Q&A



감사합니다