
개인정보 가명처리(비식별) 솔루션 비교 분석



보안뉴스

KAIST Cyber Security Research Center

Kangsik Shin, Researcher

2023. 06. 08

개인정보 수집 목적 및 활용

- 개인정보란? 이름, 주민등록번호, 휴대폰번호 등을 통해 개인을 식별할 수 있는 정보를 의미함
- 전자상거래, 고객관리, 금융거래 등 사회의 구성, 유지, 발전을 위한 필수적인 요소
- 하지만, 개인정보 불법 수집과 오남용 등으로 스팸 문자 및 불필요 광고가 증가
- 4차 산업혁명 시대로 발전함에 따라 개인정보의 범위와 영역이 더욱 확장
- 개인정보와 개인정보보호 대한 **활용과 수요**는 더욱 **증가**할 것으로 **예상**



<공공기관(좌) 및 민간 기업(우) 개인정보 수집 목적 (출처 : 개인정보보호위원회, 2022)>

개인정보 데이터의 종류

문자

- 일반 회원 정보
- 의료 정보
- 카드거래내역
- 온라인 쇼핑몰 등



음성

- 통화 내용
- 일반 상담
- 블랙박스 내 음성
- AI 음성 비서



이미지

- 일반 사진
- 스캔 된 문서
- 신분증 사본
- SNS



영상

- CCTV
- 블랙박스 영상
- 기타



국내 개인정보 제도화

1995

신용정보의 이용 및 보호에 관한 법률

신용정보의 효율적인 활용과
개인신용정보의 보호라는
두가지 목적을 위해 제정

2011

개인정보보호법

개인의 자유와 권리를
보호하고 나아가 개인의
존엄과 가치를 구현함을
목적으로 함

데이터 활용
근거 마련

2001

정보통신망 이용촉진 및 정보보호 등에 관한 법률

개인정보 보호에 대한 조항을
포함시켜 정보통신망을
이용하여 영리를 추구하는
사업자들을 대상으로 규제

2020

데이터 3법

개인정보보호법,
정보통신망법,
신용정보보호법 등의 동시
개정을 통해 개인정보
보호체계 일원화

개인정보 활용의 시작

유럽연합 개인 정보 보호법

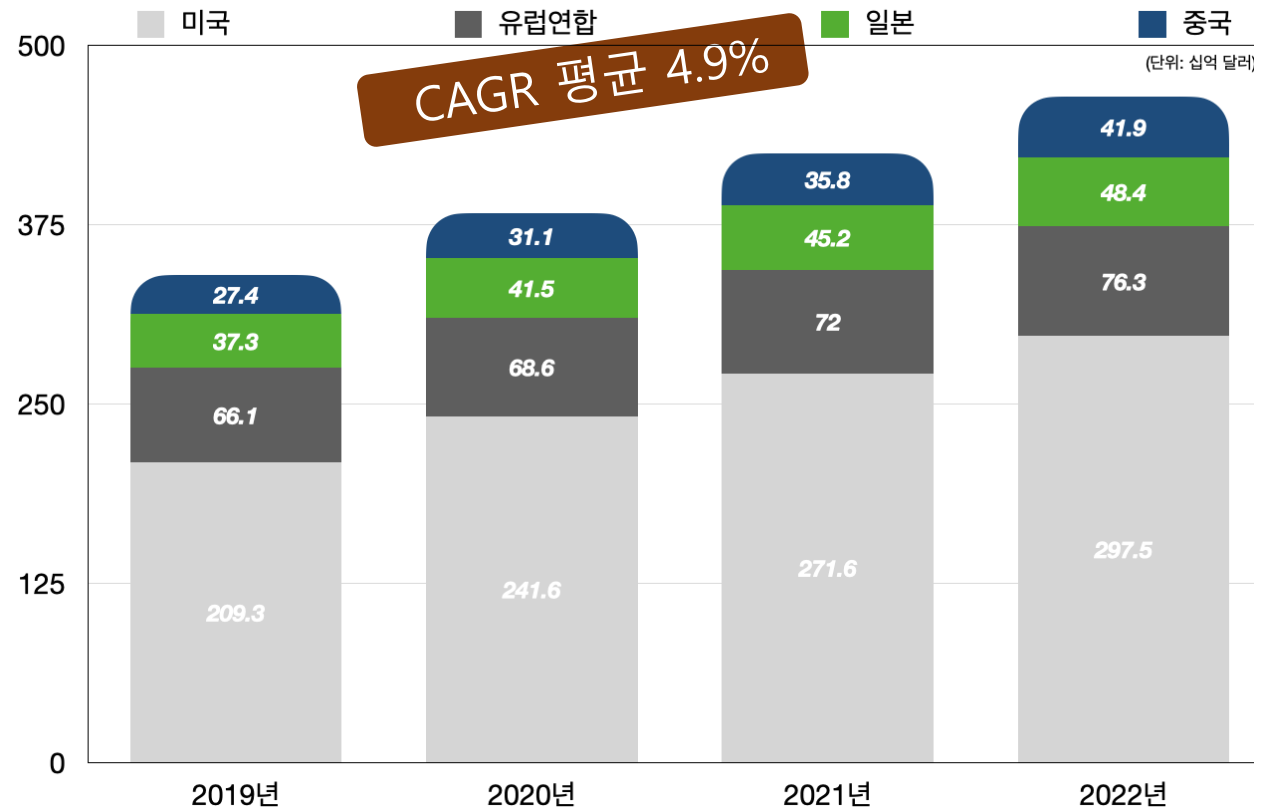
EU GDPR 발효 '18.05.25
(General Data Protection Regulation)

관련조항

- 제5조 GDPR. 개인정보 처리 원칙
 - 제6조 GDPR. 처리의 적법성
 - 제89조 GDPR. 공익적 기록보존 목적, 과학적 또는 역사적 연구 목적, 또는 통계적 목적을 위한 처리와 관련한 안전조치 및 적용의 일부 제외
- GDPR 조항, 개인정보 활용 가이드라인 정립
 - 개인 데이터를 활용한 시장성 확대
 - 개인정보 활용을 위한 **가명처리 솔루션** 필요성 대두



데이터 3법



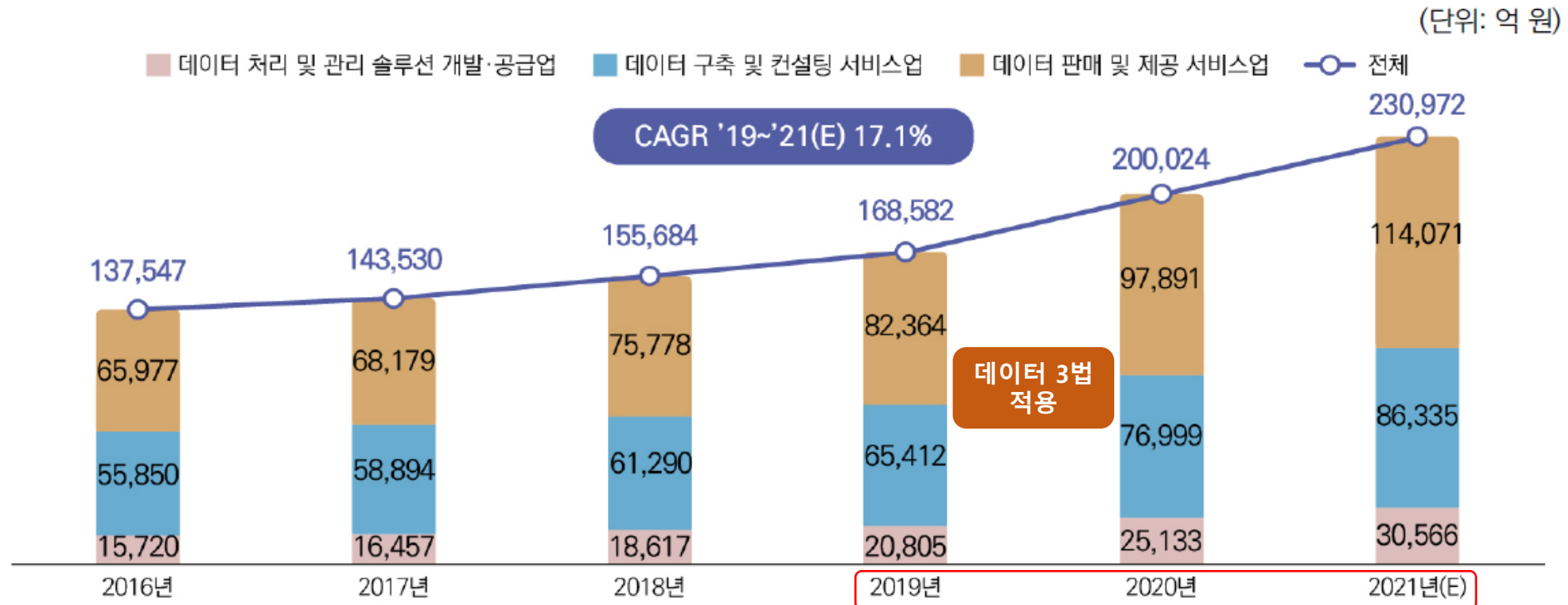
• 2019-2021: European Data Market Study(DC, 2022년 2월)

*CAGR : Compound Annual Growth Rate

글로벌 데이터 시장 규모

국내 데이터 산업 시장 현황

- 전 세계적으로 데이터는 디지털 경제를 이끄는 미래 **핵심 동력원**
- AI 기술 및 새로운 ICT 인프라와 접목되어 지속적인 성장세
- 국내 데이터 산업시장 규모의 경우 2019년부터 2021년까지 **3년 간 연평균 성장률 17.1%**
- 국내 데이터산업 시장은 **2027년까지 약 47조 원**이 넘어설 것으로 예상 됨

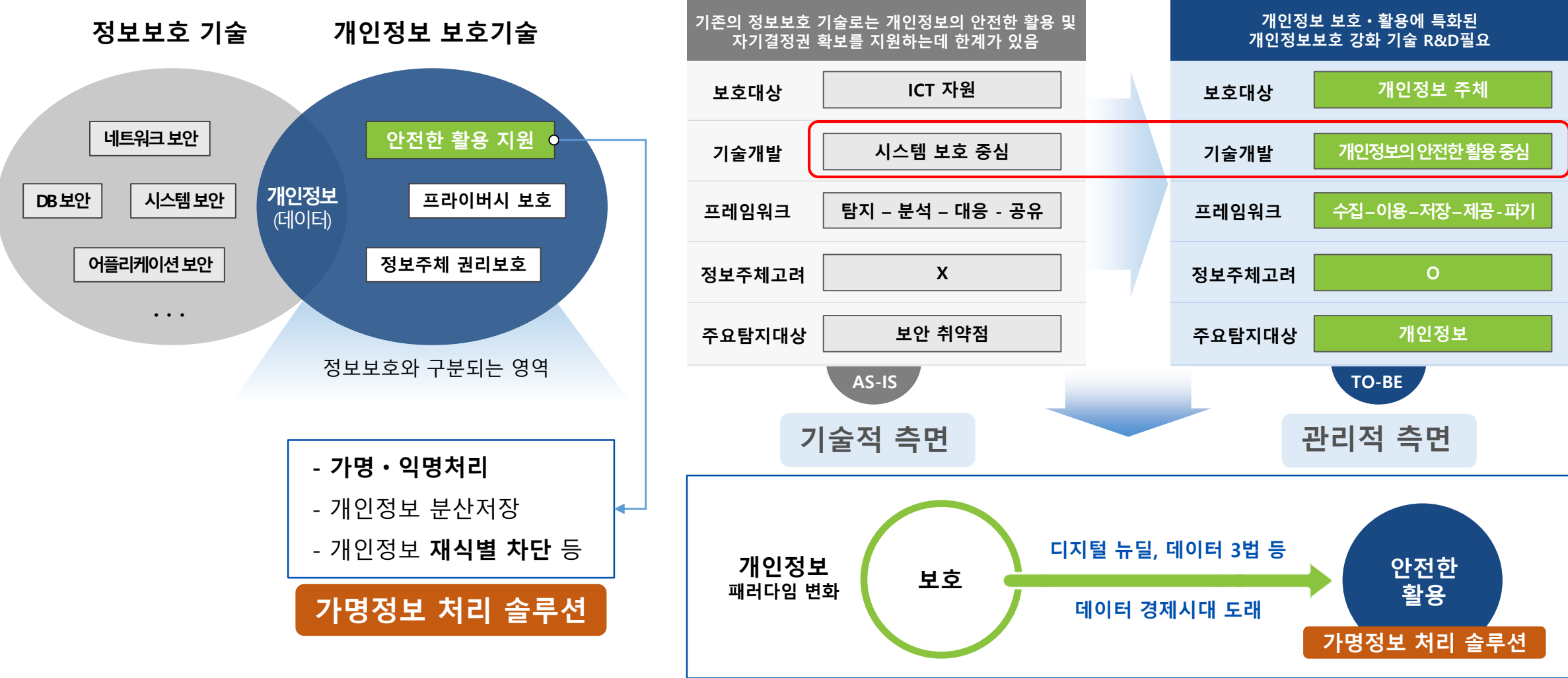


<국내 데이터 산업시장 규모, 출처 : 한국데이터산업진흥원, 2022>

*CAGR : Compound Annual Growth Rate

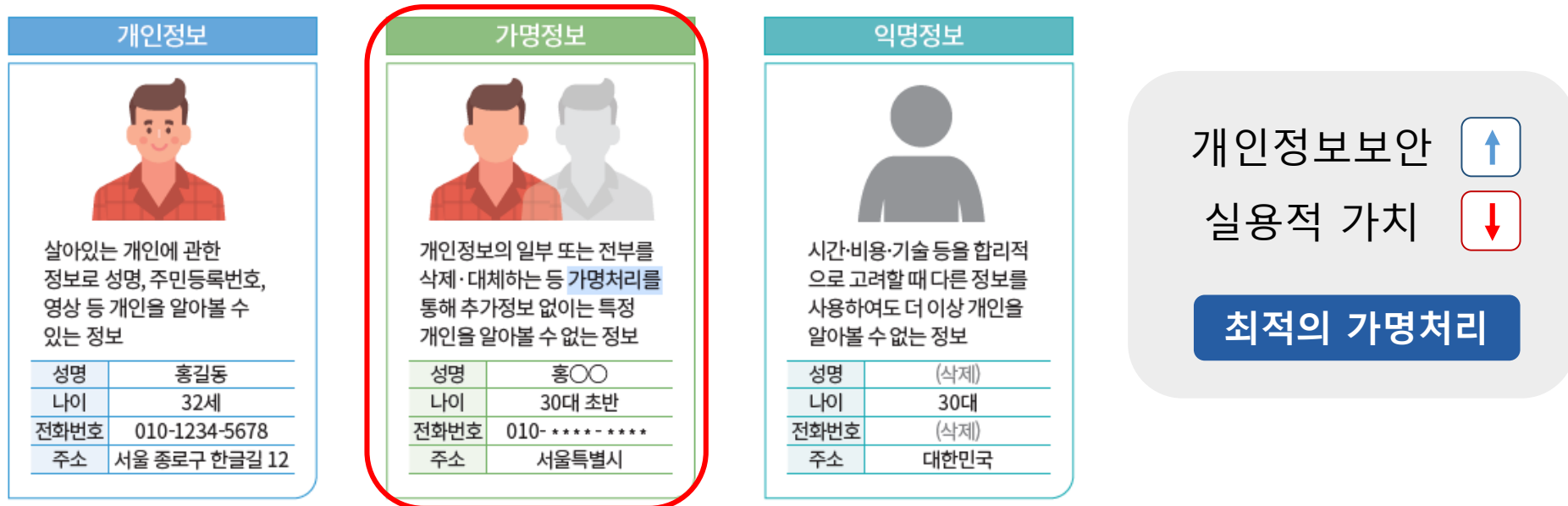
개인정보 보호는 어디에? 어떻게?

출처 : KISA



가명처리는 무엇이고 왜 하는걸까?

- 가명처리란? 비식별화 라고도 하며 누군가의 정체성이 공개되지 않도록 하는 처리 과정을 의미함
- 4차 산업혁명 시대 신성장 동력인 '데이터' 활용에 대한 시대적 요구 반영 -> 데이터3법
- 디지털 전환 가속화, 인공지능 기술 및 적용으로 개인 정보 사용 범위 확대
- 통계작성, 과학적 연구, 공익적 기록보존 등을 위한 목적으로 활용
- 특정 개인을 식별할 수 없도록 가명처리 이후 동의 없이 데이터 사용 가능



가명처리를 올바르게 못한다면...

- 인터넷쇼핑몰, 통신사, 요식업 등 홈페이지 가입 및 관리 목적의 수집된 개인정보
- 가명처리를 절차에 맞게 진행 하지 않으면, **개인정보 유출**로 이어 질 수 있음
- 유출 된 개인정보는 **보이스피싱, 스팸광고 등 범죄에 사용**되어 2차 피해로 이어짐

맥도날드, 고객 487만 명 개인정보 유출..과징금·과태

크 701

[단독] LGU+ 개인정보유출, 18만명?...해커,
"3000만건 데이터 보유"

16:32 | 수정 2023-03-22 16:33

1 1 1 1 1

개인정보위, 쿠팡 관련 개인정보 46만건 유출사고 확인

3627

좋아요 8개 | 입력: 2023-03-21 13:07



최근 다크웹 해킹포럼에 물품 구매 기록 회원 개인정보 게시됐다는 의혹
유출된 개인정보 출처, 경위, 규모 등 검토해 관련자 확인 후 법 위반 착수 예정

[보안뉴스 김영명 기자] 개인정보보호위원회(위원장 고학수, 이하 개인정보위)는 쿠팡과 관련된 것으로 보이는 개인정보 유출 사고에 대해 사실 여부를 확인 중이라고 밝혔다.

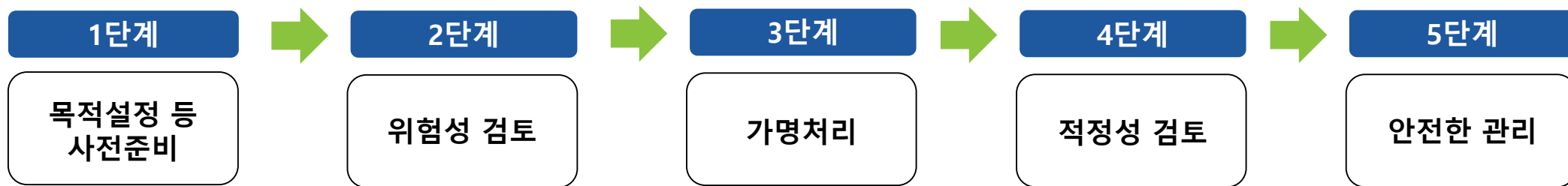
최근 쿠팡에서 물품을 구매한 기록이 있는 사람들의 개인정

coupang



가명정보 처리 솔루션이란?

- 개인정보보호위원회에서 배포하는 “가명정보 처리 가이드라인”에 준수하는 가명정보 처리 절차 기술
- 개인정보의 가명처리를 단계별 절차에 따라 처리를 도와주는 솔루션
- (1단계) 가명정보 처리 목적을 명확히 설정하고 가명정보 처리 목적의 적합성 검토 및 필요 서류 작성
- (2단계) 가명처리 대상 데이터의 식별 위험성 분석 및 평가
- (3단계) 식별 위험성 검토 결과를 기반으로 가명정보마다 적절한 가명 처리 기술 적용
- (4단계) 1~3단계의 가명처리 결과에 대한 적정성 검토(검토위원 선정 등)
- (5단계) 가명 처리된 데이터에 대한 기술/관리/물리적 안전 조치 등 사후 관리 이행



<개인정보의 가명 처리 단계별 절차도>

가명처리 단계 및 설명(1/3)

1 목적 설정 및 사전 준비

- 명확한 **가명처리의 목적**에 대한 설정(통계, 연구 등)
- 가명정보 처리에 대한 **안전조치 및 필요 서류 작성**
- 가명처리 목적에 대한 **적합성 검토 및 승인 절차**

2 위험성 검토

- 데이터 자체의 **식별 위험성 요소 확인**
- 식별정보, 식별가능정보, 특이정보, 재식별시 영향도 등을 검토
- **처리 환경 식별 위험 요소 확인**(장소, 방법, 활용 목적 등)

가명처리 단계 및 설명(2/3)

3 가명처리

- 개인정보처리자는 개인정보로부터 식별 위험성 검토 결과 분석
- 가명정보의 활용 목적 달성에 필요한 가명처리 방법 및 수준 정의
- 데이터 항목별 가명처리 계획하고 가명처리를 실행함
 - 이름 : 마스크 (예 - 홍길동 > 홍**)
 - 주민등록번호 : 삭제 (예 - 데이터 삭제)
 - 나이 : 내림(예 - 34세 > 30세)

정형 데이터			반정형 데이터	
소유자명	연락처			
김철수	010-1234-5678	서울특별시 동대문구	공동 주택	32평형
홍길동	010-4321-8765	대전광역시 서구	공동 주거 사무실	18평형
이영희	010-	대전광역시 서구	단층 주택	30평형

비정형 데이터

보증금 2천5백만원에 월 75만원
전세 팔천백이십오만원
2년계약 전세 일억이천오백

가명 처리 기법

데이터 암호화 데이터 범주화 불필요 데이터 삭제 마스킹 등							
소유자명	연락처	주택구분	주소	전세(천원)	보증금(천원)	월세(천원)	전용면적
김**	010-1234-5678	아파트	서울특별시 동작구 사당동 214-1	-	25,000	750	104.00
홍**	010-4321-8765	오피스텔	대전광역시 서구 둔산동 432	81,250	-	-	56.45
이**	010-1111	주택	세종특별시 아름동 15-1	125,000	-	-	100

식별가능정보

식별정보

데이터 가명처리

가명처리 단계 및 설명(3/3)

4 걱정성 검토

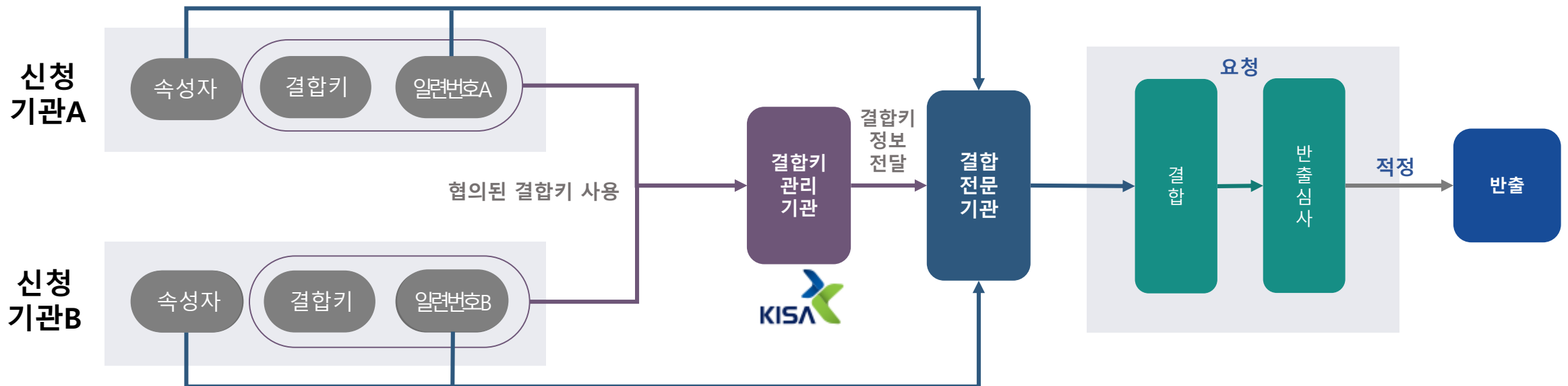
- 가명처리가 걱정하게 수행되었는지 확인
- **가명처리 결과가 가명정보의 처리 목적을 달성에 적절한지 등 검토**
- 검토는 내부 인원 또는 외부 전문가를 통해 수행 가능
- 재식별의 가능성이 있는 경우 절차를 재 수행하거나 부분적으로 추가 가명처리 수행

5 안전한 관리(모니터링)

- 가명정보는 법에 따라 기술적/관리적/물리적 안전조치 등 사후관리를 이행
- 가명정보 처리 이후에도 특정 개인이 식별 되는 경우 처리중지/회수/파기 등 조치 수행
- 가명처리 된 정보에 대한 관리를 내부 관리 계획에 따라 모니터링
- 가명정보 처리 관련 기록 작성 및 보관

가명정보 데이터를 활용한 결합

- 가명처리 솔루션 중요 기능, (외부결합) 결합전문기관에서만 사용하는 특징으로 평가항목 제외
- 서로 다른 개인정보처리자가 보유한 개인정보를 가명처리 후 결합을 통해 가치 있는 데이터 생성
- 가명정보 결합 과정에서 담당자가 특정 개인을 식별할 수 있어 신뢰받는 **제3의 기관(결합전문기관)** 필요
- 개인을 식별할 수 없는 **공통된 결합 키**를 통해 **결합 전문기관**으로 부터 **결합** 수행
- 결합 수행 후 **결합 된 데이터가 개인을 특정할 수 없는지** 심사 후 **반출 및 활용 가능**



가명정보 결합 활용 사례 (1)

친환경차 충전 인프라 수요예측 및 최적입지 선정

- (원인) 전기차 충전시설 설치사업을 추진하는 지자체는 시민 거주지 정보와 차량정보만 보유
- (문제) 실질적으로 전기차 충전시설 설치가 필요한 최적의 입지를 분석하는 데 한계가 있음
- (결과) 거주자와 근무자가 많은 아파트, 관공서 중심의 충전소 설치가 불가피함
- (방법) 친환경차를 보유한 주민의 차량정보와 차량의 이동정보를 결합, 최적의 충전소 입지 선정이 가능
- (활용1) 역세권, 대형병원 일대는 전기차 충전시설이 추가로 설치될 필요
- (활용2) 음식점, 카페 밀집 지역은 차량 방문 빈도가 높고 체류시간이 짧아 급속 전기차 충전소 설치 필요

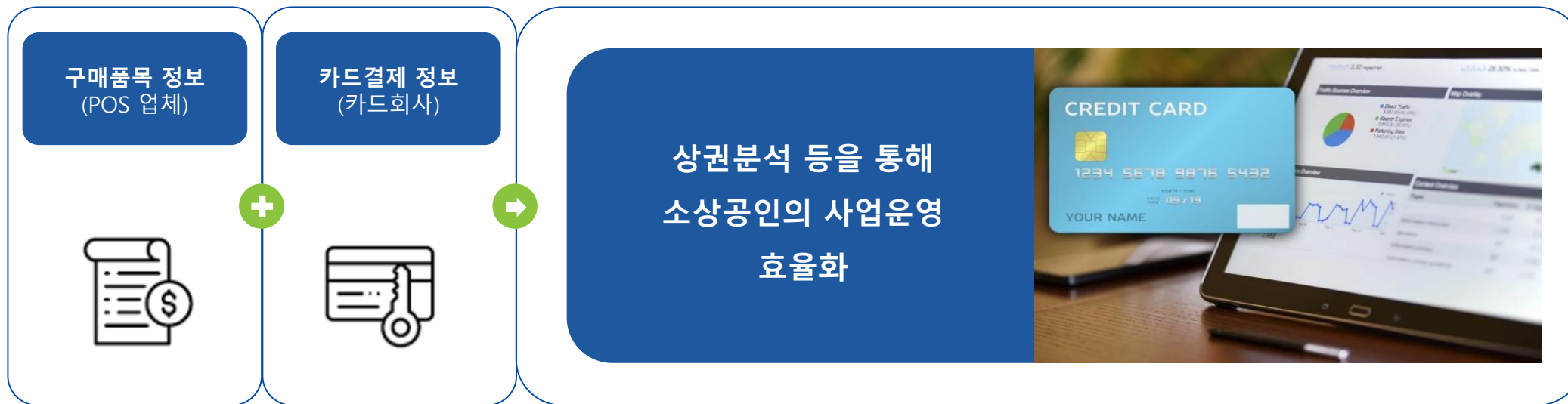


출처 : 개인정보보호위원회

가명정보 결합 활용 사례 (2)

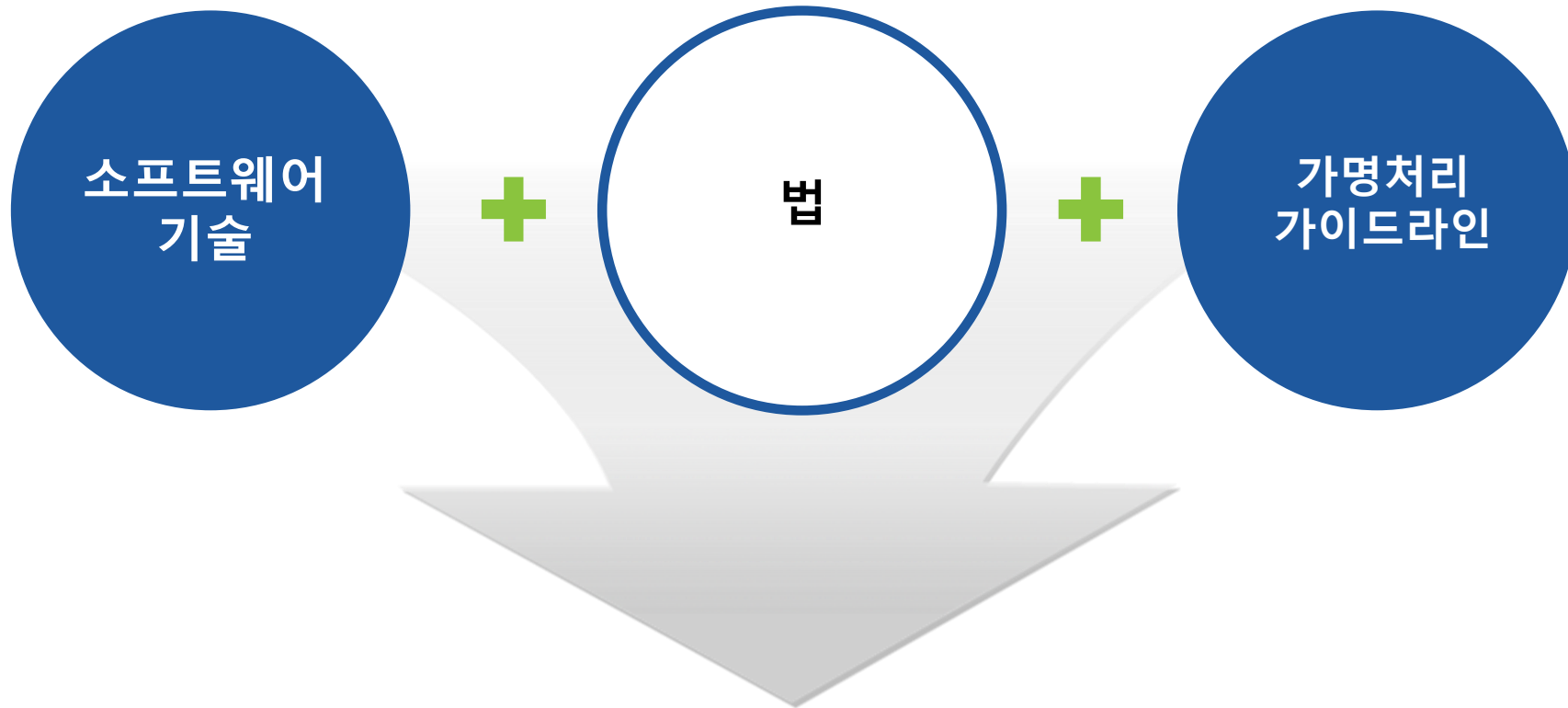
소상공인을 위한 맞춤형 고객 정보 제공

- (목적) POS기를 사용중인 10만 소상공인들의 메뉴 및 업종 표준화 등을 통한 정량적 분석
- (방법) POS기 내 영수증 정보와 카드사가 보유한 결제 정보를 결합, 요식업 소비에 관한 분석
- (분석결과) 메뉴 선호도를 파악, 주변 상권 정보를 분석해 고객-상권에 대한 메뉴 전략 수립 가능
- (활용계획1) 고객의 특성에 따라 선호 하는 메뉴와 비 선호 메뉴를 파악해 가게 마다 메뉴 전략 수립
- (활용계획2) 주기적인 반복 결합을 통해 맞춤형 고객 기반으로 소상공인들의 사업체 운영을 도움



<출처 : 개인정보보호위원회 & 픽사베이>

그래서 가명처리 솔루션이란?



법을 준수 하고, 개인정보 담당자(기업)가 개인정보를 잘 활용 할 수 있도록
가명처리 가이드라인 을 기반으로 가명처리 를 도와 주는 도구

가명처리 솔루션 평가 절차 및 일정

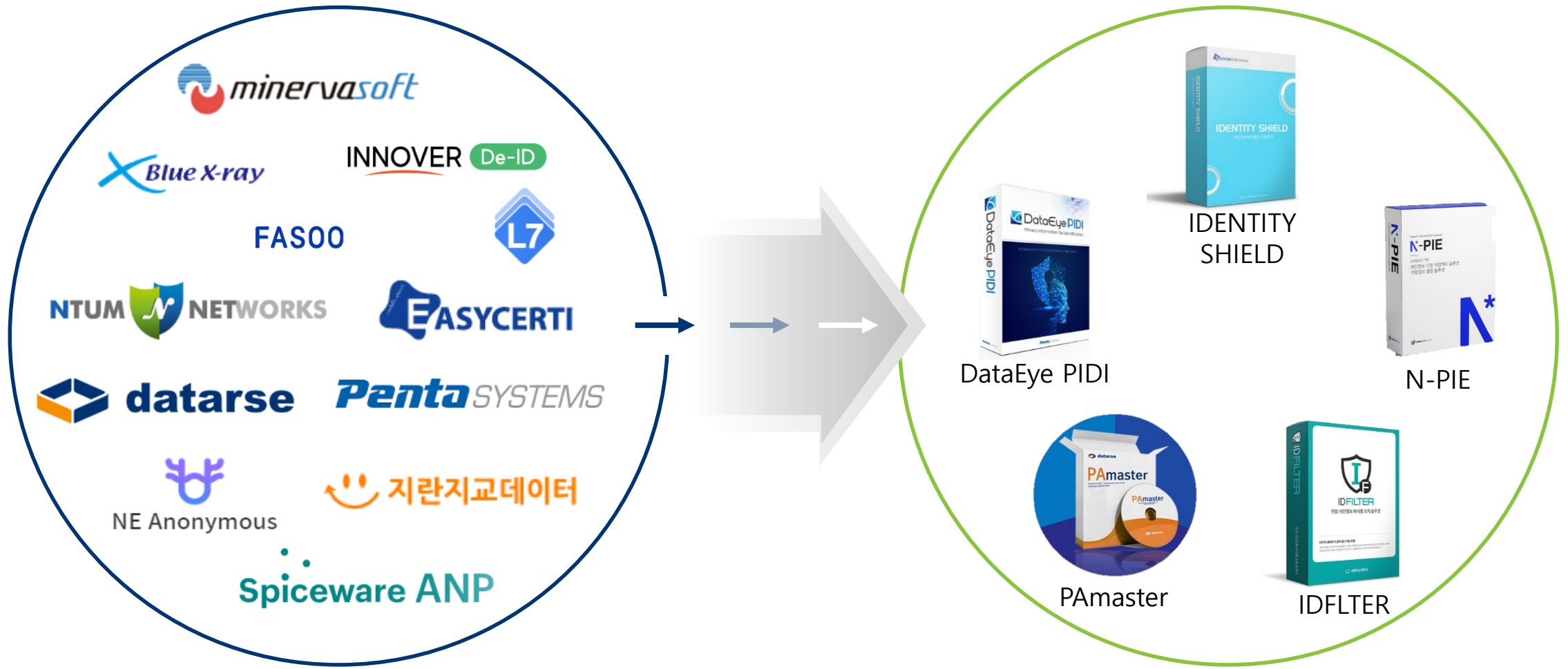


공정성을 위한 동일 스펙
CPU : Intel Xeon 8 Core
(2.1Ghz)
MEM : 32GB
Storage : SSD 1TB * 2EA
OS : CentOS 7.9



가명처리 솔루션 대상 선정

점유율 + 설문조사 인지도 + 의지



가명처리 솔루션 평가 항목 정의

- 개인정보보호위원회에서 발행한 “가명정보 처리 가이드라인(2022.04)” 분석 후 평가 항목 참고
- 외부 결합 및 반출 관련 평가는 항목에서 제외 후 약 120여개의 평가 항목을 수립
- 가이드라인에서 제시하는 정량적 평가 항목(O), UI/UX 등 편의기능(X), 제한 시간 : 39개 간소화

중분류	소분류
사전준비	가이드라인에 정의되어 있는 관련문서(위탁문서, 동의서 등) 업로드가 가능한가?
위험성검토	원본 데이터에 대한 속성 확인(문자, 숫자 등) 기능이 있는가?
	원본 데이터에 대한 미리 보기 기능이 있는가?
	속성 별 위험도 분석 (재식별 위험도 / 유사식별자 / 미국 의료정보보호법 등) 기능 제공
	위험분석시 특이 정보(전체적인 패턴에서 벗어나는 극단값) 확인이 가능한가?
가명처리 및 가명정보의 처리	데이터에 따른 가명처리 방법에 대한 추천 기능이 있는가?
	대상항목별 데이터 분석 현황 통계자료(그래프, 시각화 등) 출력이 가능한가?
	속성별 (식별정보, 식별가능정보, 특이정보, 일반정보) 위험도 유형 분석이 가능한가?
	총계처리 기능을 지원하는가? (평균값, 중앙값, 최댓값, 최솟값, 총계, 최빈값)
	가명정보에 대한 가명처리 방법 중 암호화 기능을 지원하는가?
	가명정보에 대한 가명화 처리 방법 중 매핑 테이블, 토큰 기능이 있는가?
	가명정보에 대한 마스킹(반전기능 포함), 부분삭제, 컬럼삭제와 같은 삭제 처리 방법이 있는가?
	가명정보에 대한 상하단 코딩, 라운딩 등 일반화 처리 방법이 있는가?
	가명정보에 대한 잡음 추가와 같은 무작위 처리 방법이 있는가?
	가명처리 관련 업무 담당자를 분리할 수 있는가?
정확성	100MB 가명처리 정확도 평가 (컬럼 4개, 처리방법 3개 이상)
	1GB 가명처리 정확도 평가 (컬럼 4개, 처리방법 3개 이상)
	가명정보 처리 목적에 해당하는지, 처리목적을 작성 또는 선택 할 수 있는가?
	처리 환경의 식별 위험성에 대한 점검할 수 있는가?
	적정성에 대한 기초자료 명세서 생성 또는 등록 기능이 있는가?
	개인정보 보호 수준 분석(민감, 식별 가능 정보 등) 가능 여부를 확인 할 수 있는가?

평가를 위한 테스트 데이터 생성

- 개인정보보호위원회 가명정보 처리 가이드라인(2022.04)에서 제시된 **개인 식별 가능성이 높은 정보 활용**
- 식별정보(성명, 주민등록번호 등) 및 식별가능정보(성별, 연령 등) 일부 항목
- 마스킹, 삭제, 범주화 등의 가명처리 기술을 적용할 수 있는 항목을 선정
- 보건의료, 교육, 공공분야 등의 복합적인 분야에서 활용가능한 항목이 포함된 가상 데이터 생성 도구 개발
- 공정성 있는 평가 진행을 위해 각 기업 의견을 수렴, *100MB 데이터와 1GB 데이터를 생성 함

*1MB당 약 5,000건의 데이터 수

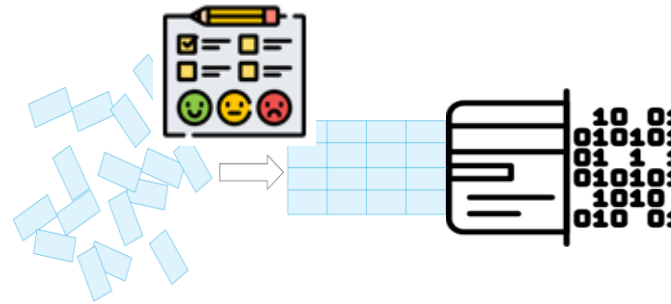


예시 개인 식별 가능성이 높은 정보

▶ 식별정보: 성명, 고유식별정보(주민등록번호, 여권번호, 외국인등록번호, 운전면허번호), (개인)휴대전화번호, (개인)전자우편주소, 의료기록번호, 건강보험번호 등 식별을 목적으로 생성된 정보

▶ 식별가능정보: 성별, 연령(나이), 거주 지역, 국적, 직업, 위치정보 등 개인정보처리자의 입장에서 개인을 알아볼 수 있는* 정보

* 개인을 '알아볼 수 있는'는 해당 정보를 처리하는 자(정보의 제공 관계에 있어서는 제공받는 자를 포함)를 기준으로 판단하여야 함



순번	이름	성별	주소	주민등록번호	나이	성지	종교	카드번호	은행	직업	기업	연락처	비고
1	김지훈	남	경주광역시 경동구 서초대거리	310117-1188596	92	163	49	127200000	5300000	대학원생	이성	062-101-3223	O
2	이준석	남	경주광역시 서구 압곡영거리	590609-1349686	64	180	53	109700000	11800000	대졸	특정 가중관리 기계 조직생	051-324-5456	A
3	조윤서	여	경상남도 사천시 도산대거리	420004-2167423	81	206	68	129500000	11300000	고졸	가족 사육 종사원	041-947-9055	A
4	김영희	남	경상북도 서천군 서초대길	630811-1688712	60	208	62	4620000	11800000	초졸 이상	공인 제관공	054-825-7214	A
5	정우현	남	경기도 화성시 동성로(문정면)	210425-1175530	102	177	47	147600000	5000000	대학원생	국립행정 사무원	055-267-5166	B
6	김민정	남	경상남도 창원시 마산합포구 (신정면)	460307-1368333	77	190	73	4300000	5000000	대졸	단체 사무원	052-564-5652	B
7	김성민	여	인천광역시 중구 남동로 (영제동)	310108-2613763	92	173	55	9630000	2700000	대졸	건설(에너지 기술자) 및 연구원	032-381-7791	B
8	홍영우	여	부산광역시 북구 대천로(남구)	880913-2874194	35	198	82	5400000	13700000	대졸	법무사 및 집행관	033-384-0588	B
9	오현우	여	서울특별시 북구 대천로(남구)	830921-2163697	40	180	72	9770000	5100000	고졸	생산 및 품질 관리 사무원	052-513-4901	B
10	김학수	여	경기도 용인시 상갈구 대천로(남구)	940321-2163691	29	173	76	4790000	5000000	중졸	생산 및 품질 관리 사무원	019-455-3026	O
11	최성민	남	서울특별시 강남구 영재대로 (서초동)	360804-1214831	87	161	59	6010000	8000000	대졸	사회복지관련 관리자	017-245-6396	B
12	강민재	남	부산광역시 관악구 연호로 (대천동)	210702-1014042	102	182	96	8930000	11700000	대졸	재능동용 이용자	043-518-1874	O
13	윤지우	남	경주광역시 동구 남동로 (서초동)	360802-1955667	84	179	78	9620000	1900000	대졸	문화유산 관련 분석가	041-955-8074	O
14	김지훈	남	경상북도 안동시 남동로	540519-2028666	69	157	70	10410000	6100000	중졸	후자 및 신문 분석가	016-158-0214	A
15	안국자	여	부산광역시 중구 남동로	650300-2028506	58	208	87	5300000	13700000	대졸	관리사	011-574-4037	A
16	윤국수	남	전라남도 목포시 개포로 (영수동)	380902-1913078	85	165	100	2500000	8800000	대졸	가족 관공청에 참여원	032-540-8444	A
17	이윤서	남	경기도 고양시 안동로	480229-1648229	75	177	97	1730000	13400000	중졸	보건의료관련 관리자	051-453-5507	B
18	박지훈	남	충청남도 홍주시 서문구 남동로	320725-2536723	91	204	89	7310000	12400000	초졸 이상	한국 우편물 및 조식	019-811-3124	O
19	박동우	여	전라북도 완주군 개포로(남동면)	370609-2464166	86	208	95	10550000	8000000	대학원생	글로벌 경영 및 책임경영	054-979-3632	A
20	지정우	남	서울특별시 서초구 개포로 (영수동)	910225-1888229	32	181	83	12000000	9200000	대졸	생물과학 사무원	019-4321-4025	A

<가명처리 데이터 생성 절차>

가명처리 솔루션 테스트 결과

- 개인정보보호위원회에서 배포한 가명정보 처리 가이드라인(2022.04)에서 제공 된 기능 위주 평가
- 가명처리 단계별 요구조건 또는 권장 기능에 대한 평가 진행
- UX/UI, 편의성 등 주관적인 평가 항목은 최대한 배제 후 평가 항목 간소화
- 2개의 대분류 6개의 중분류 39개의 소분류 항목을 통해 진행

대분류	중분류	소분류	DataEye PIDI	IDENTITY SHIELD	IDFILTER	N-PIE	PAmaster
기능성	가명정보 적정성	가명정보 처리 목적에 해당하는지, 처리목적을 작성 또는 선택 할 수 있는가?	○	○	○	○	○
		처리 환경의 식별 위험성에 대한 점검할 수 있는가?	○	○	○	○	○
		적정성에 대한 기초자료 명세서 생성 또는 등록 기능이 있는가?	○	○	○	○	○
		개인정보 보호 수준 분석(민감, 식별 가능 정보 등) 가능 여부를 확인 할 수 있는가?	○	○	○	○	○
		가명처리 된 데이터와 원본데이터와의 데이터 분포를 비교할 수 있는가?	○	○	○	○	○
		가명처리 된 데이터의 특이값(Outlier) 제거 및 관리가 가능한가?	○	○	○	○	○
		식별가능정보 동질집합(K-익명성)에 대한 분석이 가능한가?	○	○	○	○	○
		일반정보 컬럼간 연결 및 정보유출 위험을 확인할 수 있는가?	○	○	○	○	○
		추가가명 처리 기능이 있는가?	○	○	○	○	○
		위험도 측정결과 및 적정성 종합판단 및 가명처리 결과(유일값, 빈도값 변화량) 확인 가능한가?	○	○	○	○	○
		가명처리 결과 및 데이터 분석에 대한 보고서 생성이 가능한가?	○	○	○	○	○
		가명 처리 결과가 가명정보의 처리 목적을 달성하기 위해 적절한지 등 검토할 수 있는가?	○	○	○	○	○
		적정성검토에 대한 심사상태를 모니터링 할 수 있는가?	○	○	○	○	○
		가명정보처리 가이드라인(22.04 기준) 가명처리 단계별 절차에 맞게 진행되는가?	○	○	○	○	○
		검토결과 및 의견을 개인정보처리자에게 전달할 수 있는가?	○	○	○	○	○

100% 가이드라인 충족

가명처리 정확도 평가 방법 및 결과

- 생성한 가명처리 데이터 컬럼 28개 중 8개의 컬럼을 가명처리 진행
- 마스킹, 라운딩(올림), 컬럼삭제 3개의 가명처리 기술을 적용
- 개인정보 100MB(420만개), 1GB(4300만개)에 대한 정확도 테스트 진행
- 5개의 가명처리 솔루션 모두 100% 정확도로 가명처리 완료

항목	가명처리 방법
이름	마스킹(두번째 글자)
주민등록번호	컬럼삭제
카드번호	컬럼삭제
주소	컬럼삭제
나이	일반라운딩(10의자리)
키	일반라운딩(10의자리)
몸무게	일반라운딩(10의자리)
전화번호	컬럼삭제
그외	미적용



이름	성별	나이	키	몸무게	직업	혈액형	혼인여부	종교	순번	카드번호	주소	전화번호	그외
박*정	여	80	150	80	만화가 및 만화영화 작가	B	N	불교	7335	45600000			
지*주	남	30	170	80	무용가 및 안무가	A	Y	무교	215269	78800000			
양*웅	여	40	170	90	기타 의복 제조원	O	N	불교	155031	82400000			
박*호	여	60	180	80	제품 생산관련 관리자	B	Y	무교	181352	69600000			
한*자	남	100	180	60	매장계산원 및 요금정산원	AB	N	힌두교	291547	49000000			
이*진	남	30	200	60	기타 의료/복지 관련 서비스 종사원	A	N	힌두교	157479	10800000			
강*진	남	50	170	60	운송장비 조립원	AB	N	천주교	397936	102500000			
황*웅	여	100	160	70	실내장식 디자이너	O	Y	힌두교	354518	82600000			
박*자	여	60	180	60	기타 여가 및 스포츠 관련 종사원	A	N	천주교	23053	120900000			
김*식	남	70	160	50	바텐더	B	N	무교	210988	114600000			
배*식	남	70	170	60	통신 및 관련 장비 설치 및 수리원	A	Y	힌두교	152749	69800000			
강*지	여	70	180	90	청소원	O	Y	무교	260555	107400000			
노*후	여	40	150	60	고무 및 플라스틱 제품 조립원	O	Y	기독교	12821	13000000			
김*수	여	50	180	60	금속가공관련 제어장치 조작용	B	Y	천주교	210065	142600000			
박*수	남	90	200	40	노점 및 이동 판매원	AB	N	기독교	70492	33500000			
박*영	남	80	160	40	여행 및 관광통역 안내원	A	Y	천주교	341328	124100000			
신*우	여	90	200	70	방사선사	A	Y	기독교	267461	133000000			
이*민	여	60	170	70	번호사	A	Y	기독교	218328	49000000			
이*현	여	60	150	90	응급구조사	AB	Y	불교	285518	61600000			
이*영	여	60	190	90	문화/예술/디자인 및 영상관련 관리자	AB	Y	무교	182639	131200000			
김*민	남	50	190	80	정부 및 공공 행정 전문가	A	Y	힌두교	377761	139200000			
박*숙	남	70	200	50	상품기획 전문가	B	N	무교	13707	143600000			
이*옥	여	90	150	50	재단사	O	N	불교	13678	110000000			
20	임*영	남	충청북도 천안시	3/DIqEbGQrVQIwsKzjTL/Q==	64	180	90	93000000	10000000	중출	기사 및 논설위원		
21	최*현	여	충청북도 괴산군	1VTU4u76K6fA4e5iSkDwQ...	57	178	58	99000000	10800000	고졸	요리사		

PAmaster 특징점



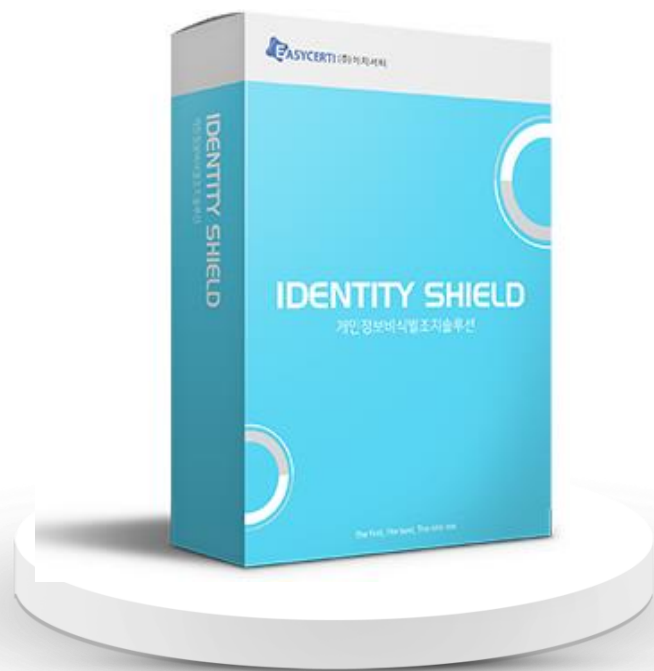
- (주)데이타스 솔루션
- 위험 지표 및 매트릭스 기반의 정량적 위험도 측정
- 위험도 측정 결과에 따른 가명처리 기술 추천
- 데이터 별 가명처리 관리 기능 제공
- 반출 심사를 위한 정량적 평가 결과 제공
- 생성부터 파기까지 가명 정보 라이프 사이클 통합 관리 기능
- 데이터 정제 및 데이터 현황 분석
- 특이 정보 식별에 따른 추가 가명 처리
- 완전 삭제 알고리즘을 통한 데이터 파기 및 관리

N-PIE 특징점



- (주)엔텀네트웍스 솔루션
- 비정형 데이터 처리 기능
- 한글파일(XML)에 대한 개인정보 가명 처리 기능
- 신청서 및 처리 결과에 대한 보고서 생성 기능
- 단일, 이중 방식의 결합 기능 제공
- 데이터 별 가명처리 관리 기능 제공
- 데이터 결합 관리 기능 제공
- 가명처리 모니터링 기능 제공
- 추출 결합 및 모의 결합 기능

IDENTITY SHIELD 특징점



- (주)이지서티 솔루션
- 비정형 데이터 내 개인정보 탐지
- 다양한 데이터베이스(4개 이상) 접속을 통한 자동 가명 처리 지원
- 클라우드(Amazon, MS Azure) 환경 구축 가능
- 30개 이상의 최다 가명/익명 처리 기술 보유
- 반출심사(3명의 심사위원/위원장 지정) 지원
- 적정성 평가에 활용 할 수 있는 통계기법 기반의 특이치 확인 기능
- 최다 데이터 결합 및 관련 기관 사용
- 대외인증 7건 이상(조달청 혁신제품, GS 1등급, V&V 성능평가 등)

IDFILTER 특징점



- (주)지란지교데이터 솔루션
- 프로젝트 별 데이터 주기 추적 가능
- 프로젝트 단계별 수행 담당자 설정 및 관리자 승인/반려 기능
- 송수신 클라이언트를 통한 암호화 송수신 프로그램
- SAS 의료 데이터 포맷 가명처리 및 결합 지원
- 데이터 대용량 파일 전송(이어올리기/이어받기)
- 데이터 송수신 시 암호/압축/무결성 검증 가능
- 결합용 대체 키 생성 기능(데이터 추적용)
- Storage, SFTP, Database 연동 기능 지원
- 과학기술정보통신부 지정, '2022년 우수정보보호 기술' 지정

DataEye PIDI 특징점



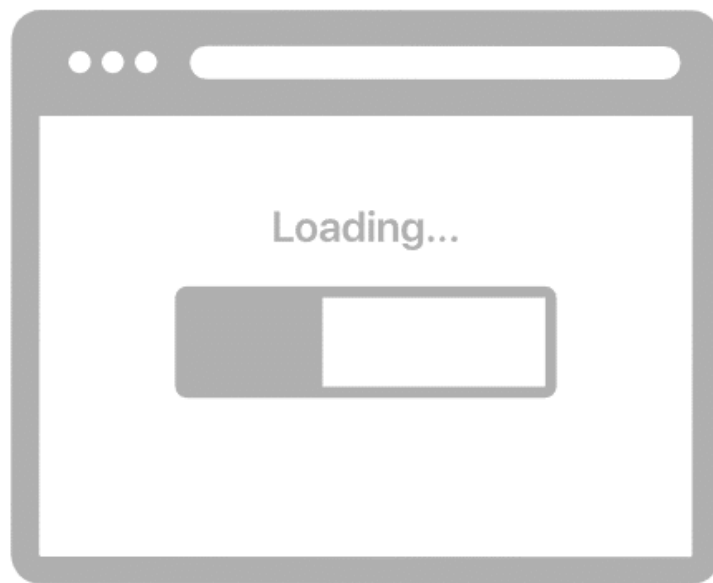
- 펜타시스템테크놀로지(주) 솔루션
- 가명처리 업무 프로세스 통합 관리 기능(Portal 제공)
- **가명처리 프로세스별 결재 기능 제공**
- **클라우드(AWS S3, Google BigQuery) 기반 연계 가능(커넥터 지원)**
- **가명처리 전/후 비교 및 차트 분석**
- 대용량 데이터셋(TB) 이상 병렬 처리 기술 보유
- 가명처리에 대한 사용자 정의 함수(기능) 제공
- 다중 가명처리 가능 및 사용자 정의 함수 100여개 제공
- 보안 컴플라이언스를 위한 보고서 제공

솔루션 보완사항

- 각 기업에서 요구하는 하드웨어 스펙 상이, 테스트시 최소사양보다 다소 부족한 하드웨어(서버) 이용
- 각 솔루션은 동일한 하드웨어 및 OS 기준으로 테스트 진행
- 100MB 데이터의 경우 모든 솔루션이 원활하게 처리 됨
- 일부 솔루션의 경우 1GB 이상의 데이터 처리시 성능 저하 및 불안정한 서비스 => 최적화 필요



가명처리



<업로드 및 가명처리 불가>



< 솔루션 서비스 불안정함 및 성능 저하 >

솔루션 보완사항

- 가명처리 기술에 대한 표준(기준)의 부재로 처리 방식 상이

<예시>

솔루션 A

원하는 위치 마스킹 가능
예) 홍*동 가능(자리수)

컬럼명 제외 후 삭제 불가
(컬럼명 보존 필요)

부분삭제 (구분자-공백)
대전광역시 유성구 => 대전광역시

솔루션 B

마스킹 특정 위치 불가/어려움
예) 홍*동 불가

컬럼명 제외 후 삭제 가능

부분삭제 가능
(구분자 입력 불가)

솔루션 보완사항

- 기능 및 목적에 따라 솔루션이 나뉘져 있어 평가 시 복잡함이 있음
- 가명처리 순서가 복잡해 다음으로 수행할 단계에 대한 안내 메시지 필요
- 결재 대상자(승인 처리자) 로그인 시 팝업이나 알림으로 결재 문서에 대한 알림 메시지 표시
- 데이터 가명처리 시 진행 사항 안내에 대한 미흡
- 가명처리 솔루션 사용 시 사용자(담당자)의 개입이 필요한 경우가 많음



<관리적 측면>

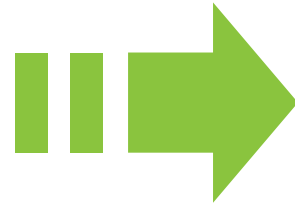


<사용자 편의성>



<담당자 입장>

가명처리
가이드라인



가명처리
솔루션

제품마다 편의성 및 다양한 특색으로 인해
기업의 니즈에 맞는 가명처리 솔루션 선택 필요

감사합니다