



데이터3법 시행에 따른  
‘개인정보 비/식/별/화 솔루션’ 소개



2021.02

# CONTENTS

## 목차

1 솔루션 개요

2 구축사례

3 구성도

4 주요기능

5 특/장점

# 1. 솔루션 개요

DataEye PIDI 솔루션은 빅데이터 환경에서 개인정보를 안전하게 활용 가능할 수 있도록 개인정보가 포함된 데이터셋을 가명화 및 익명화하여 비식별화를 수행하는 솔루션입니다.



- 제품명 : DataEye PIDI
- 용도 : 개인정보 비식별화 솔루션
- 제조사 : 펜타시스템테크놀러지(주)
- 출시년월 : 2016년 10월

- 주요기능
  - 데이터셋 사전검토
  - 비식별 조치/실행
  - 적정성 평가
  - 사후관리
  - 시스템 관리



## ▶ 솔루션 기능 및 특장점

| 비식별 조치                                                                                                                                                                                                                  | 적정성 평가(계량분석)                                                                                                                                                                                                          | 사후 관리                                                                                                                                                   | 시스템 관리                                                                                                                                                       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>• 개인정보 비식별 가이드 라인 준수</li> <li>• 데이터 가명 처리 변환</li> <li>• 특이 데이터 처리 (부문통계, 특정 범위 데이터 변환, 레코드 삭제 등...)</li> <li>• 계량분석 처리 (그룹/범주 계층 정의, 계량치 설정)</li> <li>• 검증된 암호화 기법 사용</li> </ul> | <ul style="list-style-type: none"> <li>• 비식별 조치 결과 적정성 평가 관리 (k, l, t값 평가)</li> <li>• 대용량 데이터 환경에서의 개인정보 보호 모델 처리 알고리즘 탑재</li> <li>• 평가기초 자료검토</li> <li>• 데이터셋 명세</li> <li>• 비식별화 내역</li> <li>• 비식별화 적용 기법</li> </ul> | <ul style="list-style-type: none"> <li>• 사후관리를 위한 비식별 조치</li> <li>• 결과에 대한 로그 관리 기능</li> <li>• 비식별 처리 스케줄 이력 관리</li> <li>• 설정이력/처리내역/접속기록 조회</li> </ul> | <ul style="list-style-type: none"> <li>• 순수 Web 환경의 사용자 UI 제공</li> <li>• 다양한 데이터베이스 연결 기능 제공</li> <li>• 사용자 관리, 권한 관리</li> <li>• 사용자 접속이력/사용이력 관리</li> </ul> |

## 2. 구축 사례

펜타시스템의 개인정보 비식별화 솔루션 주요 납품 실적입니다.

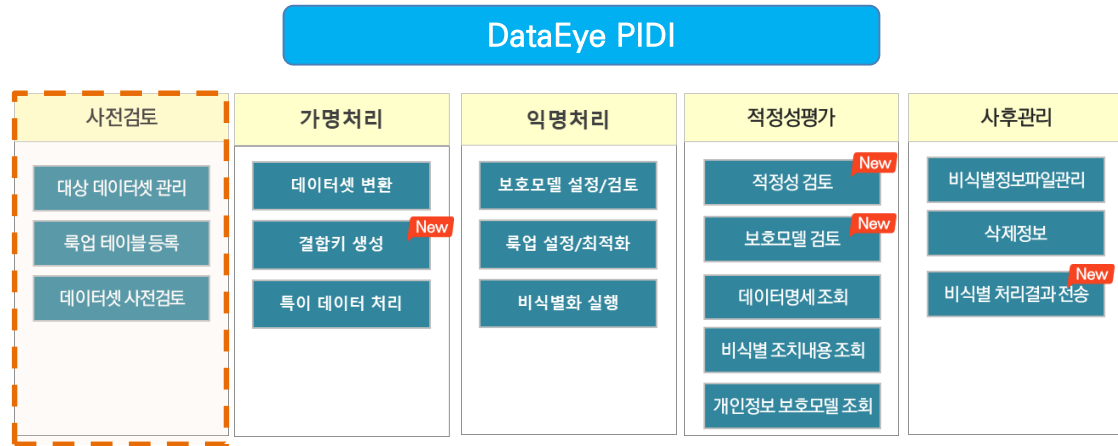
| 순번 | 사업명                        | 사업기간              | 발주처     |
|----|----------------------------|-------------------|---------|
| 1  | 데이터 비식별화 플랫폼 구축            | 2021.02 ~ 진행중     | 삼성생명    |
| 2  | 개인(신용)정보 비식별 솔루션 구축        | 2021.01 ~ 진행중     | 신한금융투자  |
| 3  | 세종시 자율주행 빅데이터 관제센터 구축 및 운영 | 2021.01 ~ 진행중     | 세종시     |
| 4  | 가명처리 솔루션 도입                | 2020.11 ~ 진행중     | KT      |
| 5  | 빅데이터 통합 플랫폼 구축             | 2020.05 ~ 2020.12 | 경찰청     |
| 6  | 빅데이터 분석 플랫폼 구축             | 2019.12 ~ 2020.03 | 우리은행    |
| 7  | 신용정보 빅데이터 시스템 구축           | 2017.09 ~ 2018.06 | 한국신용정보원 |
| 8  | 빅데이터 분석 플랫폼 구축             | 2017.05 ~ 2017.08 | 우리은행    |
| 9  | 빅데이터 플랫폼 구축                | 2017.03 ~ 2017.04 | LG유플러스  |

# 3. 구성도 - 시스템 구성도

DataEye PIDI는 데이터베이스, 파일 데이터를 개인정보 비식별 조치 가이드라인에 따른 비식별 처리의 전 과정을 시스템을 통하여 처리하도록 구성 되어 있습니다.

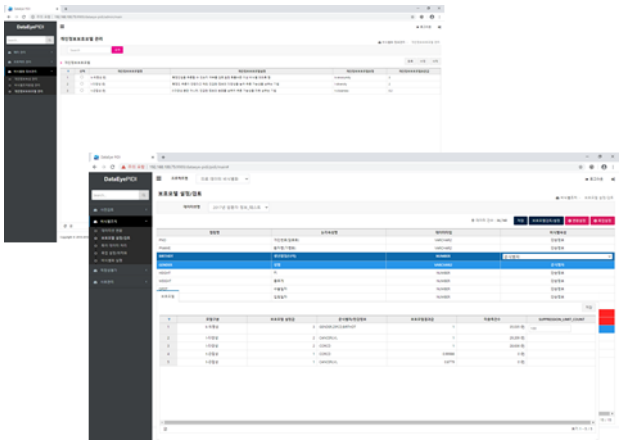


# 4. 주요기능 - 사전검토



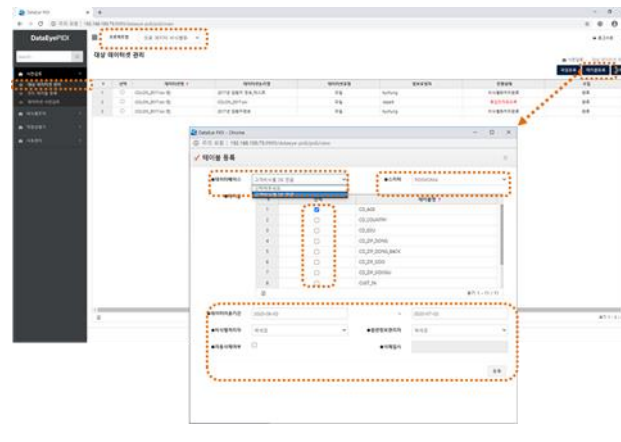
## 개인정보보호모델 관리 및 설정/검토

- 개인정보 보호 모델 관리
- 보호모델을 프로젝트에 설정 및 검토



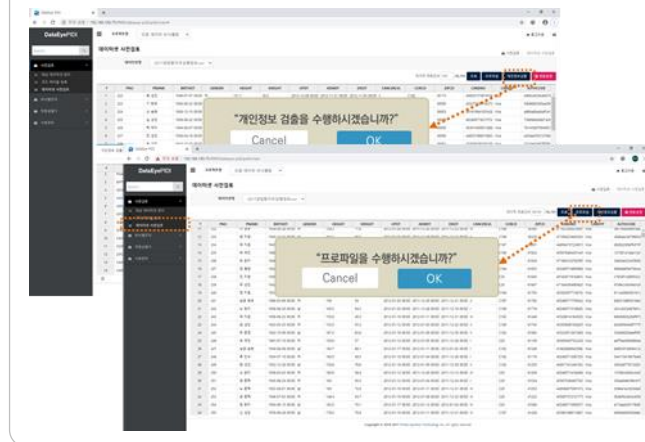
## 대상 데이터셋 관리 및 설정

- 대상 데이터셋 등록 및 삭제
- 데이터셋에 대한 속성자 정의 및 자동 분류

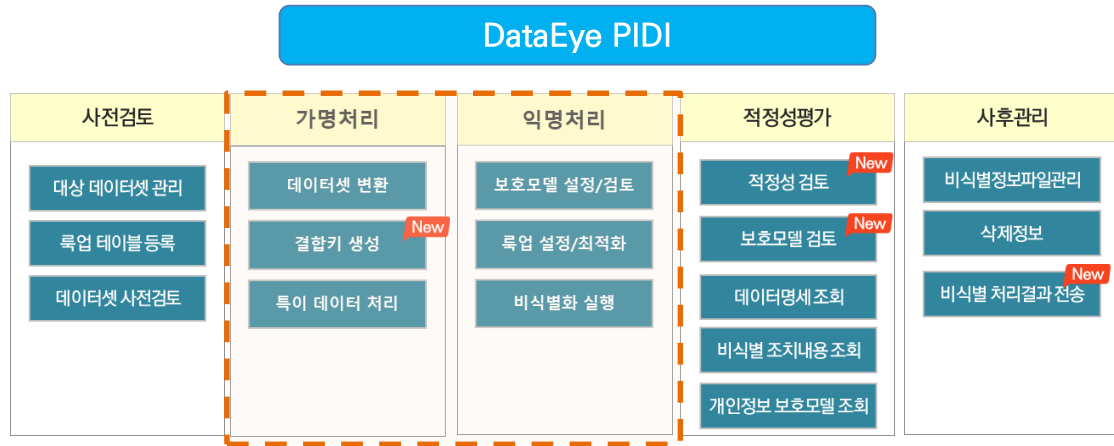


## 데이터셋 사전 검토

- 개인정보 속성 관리를 통한 개인정보 검출 작업
- 프로파일링을 통한 데이터 검증 작업

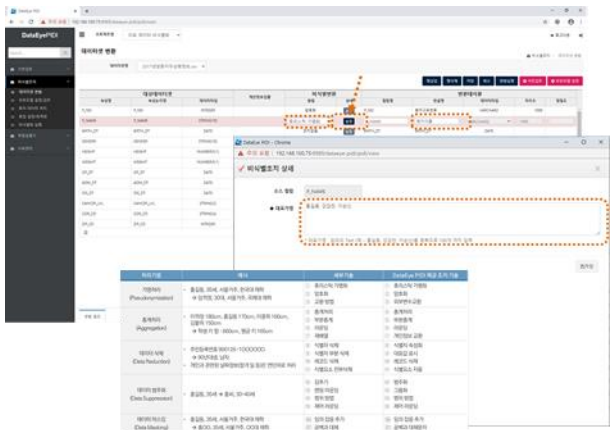


# 4. 주요기능 - 가명/익명처리



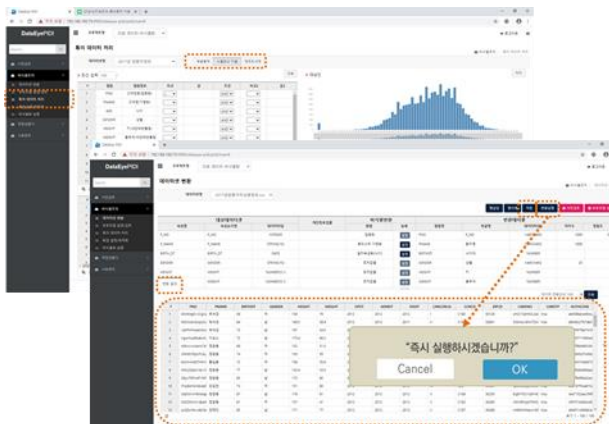
## 데이터셋 변환을 통한 가명/익명 처리

- 개인정보 비식별 가이드라인에서 제시하는 17가지 가명/익명 처리 기술에 의한 데이터셋 변환 작업



## 특이데이터 처리 및 비식별화 실행

- 특이데이터 처리를 통한 비식별 조치
- 비식별화 실행 후 데이터 확인



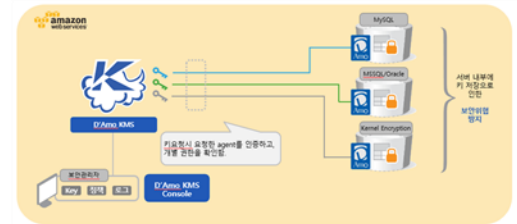
## ✓ 추가 모듈

## 암복호화 키 관리

- CC인증된 펜타시큐리티 D'Amo KMS 연동을 통한 암복호화 키 관리 기능 제공

### D'Amo KMS (Key Management System)

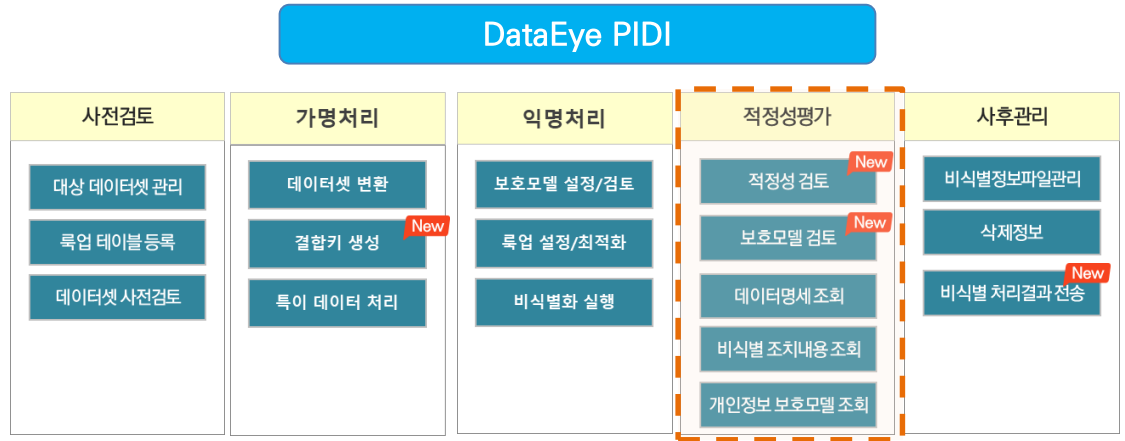
- 암호화 키를 안전한 저장소인 KMS에서 관리하여, 서버 내부에 키 저장으로 인한 보안위험을 방지.
- GUI Console로 편리하게 키를 관리를 할 수 있음



### 참고: 암호화키 관련 Compliance

- 암호키는 별도의 안전한 장소에 보관하고 접근권 부여를 최소화해야 한다. - KISA "정보보호 관리체계 (SMS)"
- 암호키는 별도의 암호화 키로 암호화되어 안전한 암호화장치에 저장해야 한다. - PCI-DSS v3.0(카드사 국제 보안 표준)
- 암호키는 DBMS와 분리된 별도의 키관리서버에 암호화하여 저장해야 한다. - 금융보안연구원 "DB암호화분석 보고서"

# 4. 주요기능 - 적정성평가



### 목표수준 설정 및 적정성 평가

- 보호모델 목표수준에 해당하는 적정성 평가 작업 수행
- 적정성 평가 결과 보호모델 조회

### 데이터셋 파기 기능 및 부적합 안내

- 데이터셋 파기 기능 제공 및 해당 프로세스에 맞게 커스터마이징 제공 가능

### 개정법령 지원 방안

- 데이터3법 개정 법령에 관한 시행령/가이드라인 준수
- DataEye PIDI 제품 업그레이드 수행

| 개정법령에 따른 제품 업그레이드 절차 |                                                                                                                                                         |
|----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| 데이터3법 법령 개정에 따른 시행령  | <ul style="list-style-type: none"> <li>✓ 데이터 3법 법령 개정에 따른 시행령 공표</li> </ul>                                                                             |
| 가명정보 처리 가이드라인 제시     | <ul style="list-style-type: none"> <li>✓ 시행령에 대한 가명정보 처리 가이드라인 세부지침 제시</li> </ul>                                                                       |
| 비식별 가이드라인 (행정안전부)    | <ul style="list-style-type: none"> <li>✓ 행정안전부의 개인정보-비식별-조치 가이드라인</li> </ul>                                                                            |
| 기능 업그레이드             | <ul style="list-style-type: none"> <li>✓ 가이드라인에 제시되는 비식별 조치방법 추가</li> <li>✓ 가이드라인에 제시되는 프라이버시 보호모델 추가 적용</li> <li>✓ DataEye PIDI 제품 업그레이드 공표</li> </ul> |



# 4. 주요기능 - 사후관리

## DataEye PID



### 가명/비식별 처리 및 삭제관리

- 가명/비식별 처리 데이터셋 관리
- 삭제 정보 관리 및 증적관련 로그 관리

1 관계 법령에 따른 증적 처리 (사후관리)비식별정보파일관리, 삭제정보 관리에서 관리

2 관계 법령에 따른 3년간 증적 처리

정보제공일시  
비식별화/익명  
수정기간  
수정일  
정보제공자  
정보위탁유무  
3차제공 또는 위탁유무

3년간 보존

### 사용자/계정 관리

- 사용자 관리 및 계정관리
- 메뉴에 해당하는 접근 권한 관리

### 시스템 모니터링 및 통계기능

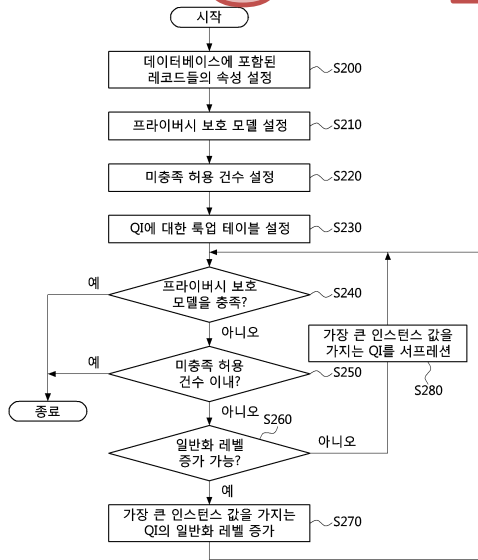
- 시스템 접속 로그 및 사용 로그
- 사용현황 통계 기능 제공

# 5. 특/장점 - 비식별화 특허 기술 활용

제안사가 보유한 개인정보 **비식별화 특허 기술을 활용**하여 정보 손실을 최소화 통하여 비식별 작업을 수행합니다.



**정보 손실을 최소화 알고리즘**



기존 방식의 Suppression

| Rec-No | QI-1 | QI-2 | QI-3 | QI-4 |
|--------|------|------|------|------|
| 1      |      |      | *    |      |
| 2      |      |      | *    |      |
| 3      |      |      | *    |      |
| 4      |      |      | *    |      |
| 중략     |      |      | *    |      |
| 5650   |      |      | *    |      |
| 5651   | *    | *    | *    | *    |
| 5652   | *    | *    | *    | *    |
| 5653   | *    | *    | *    | *    |
| 5654   | *    | *    | *    | *    |
| 5655   | *    | *    | *    | *    |
| 중략     | *    | *    | *    | *    |
| 5690   | *    | *    | *    | *    |
| 5691   | *    | *    | *    | *    |
| 5692   | *    | *    | *    | *    |

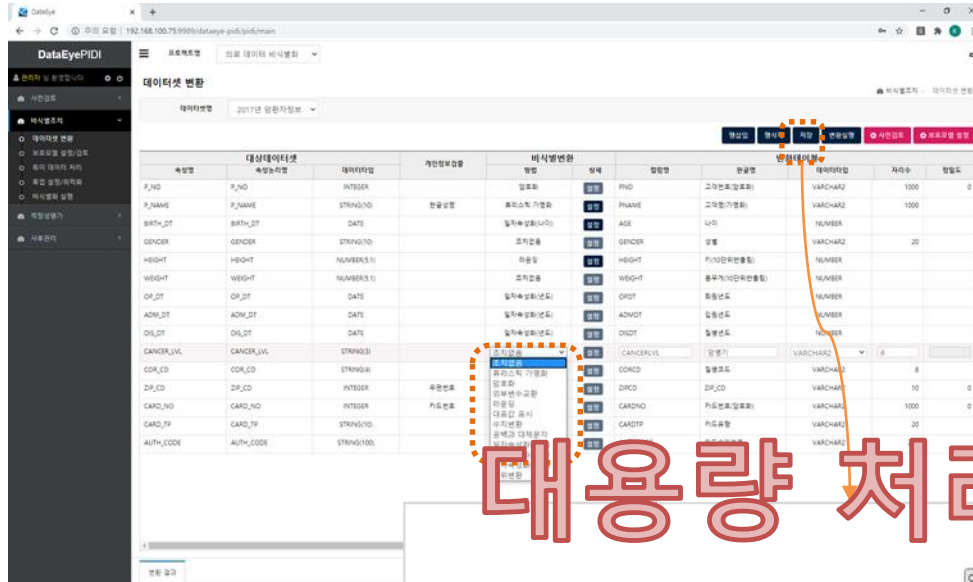
특허 방식의 Suppression

| Rec-No | QI-1 | QI-2 | QI-3 | QI-4 |
|--------|------|------|------|------|
| 1      |      |      |      |      |
| 2      |      |      |      |      |
| 3      |      |      |      |      |
| 4      |      |      |      |      |
| 중략     |      |      |      |      |
| 5650   |      |      |      |      |
| 5651   | *    | *    | *    | *    |
| 5652   | *    | *    | *    | *    |
| 5653   | *    | *    | *    | *    |
| 5654   | *    | *    | *    | *    |
| 5655   | *    | *    | *    | *    |
| 중략     | *    | *    | *    | *    |
| 5690   |      |      | *    | *    |
| 5691   |      | *    | *    | *    |
| 5692   | *    |      | *    | *    |

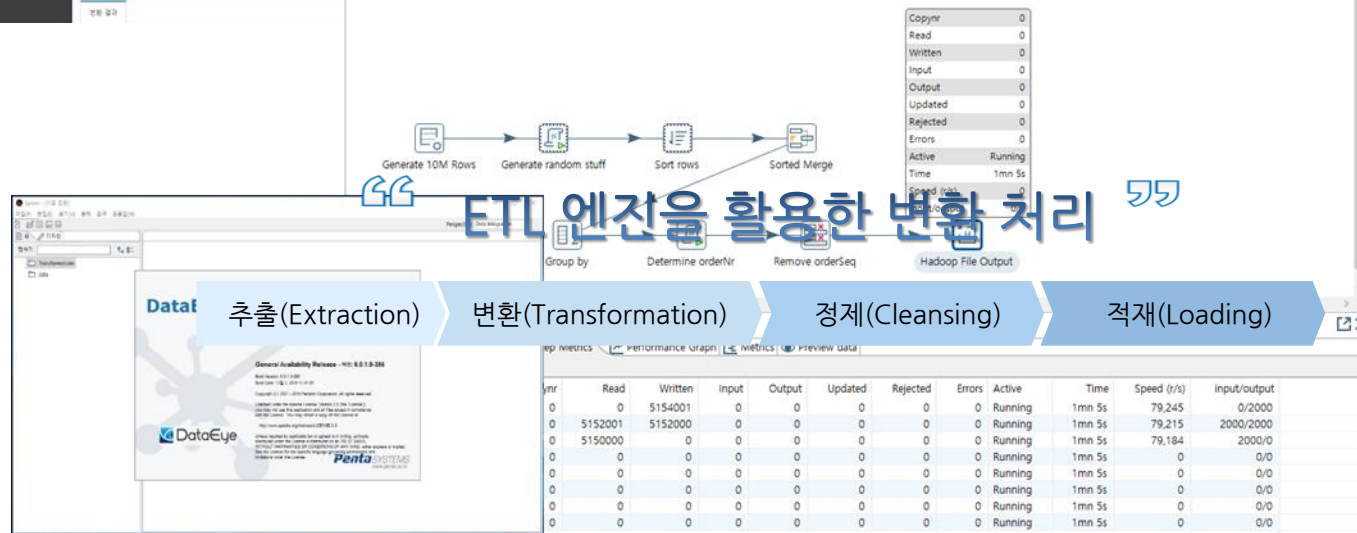
- 특허 방식의 장점
- 1 데이터 손실을 최소화
  - 2 QI(준식별자)의 개수에 관계없이 처리가 가능
  - 3 비식별화 처리 성능 극대화

# 5. 특/장점 - 대용량 처리 아키텍처

대용량 데이터 가명 처리를 위한 ETL 엔진을 활용한 성능 개선

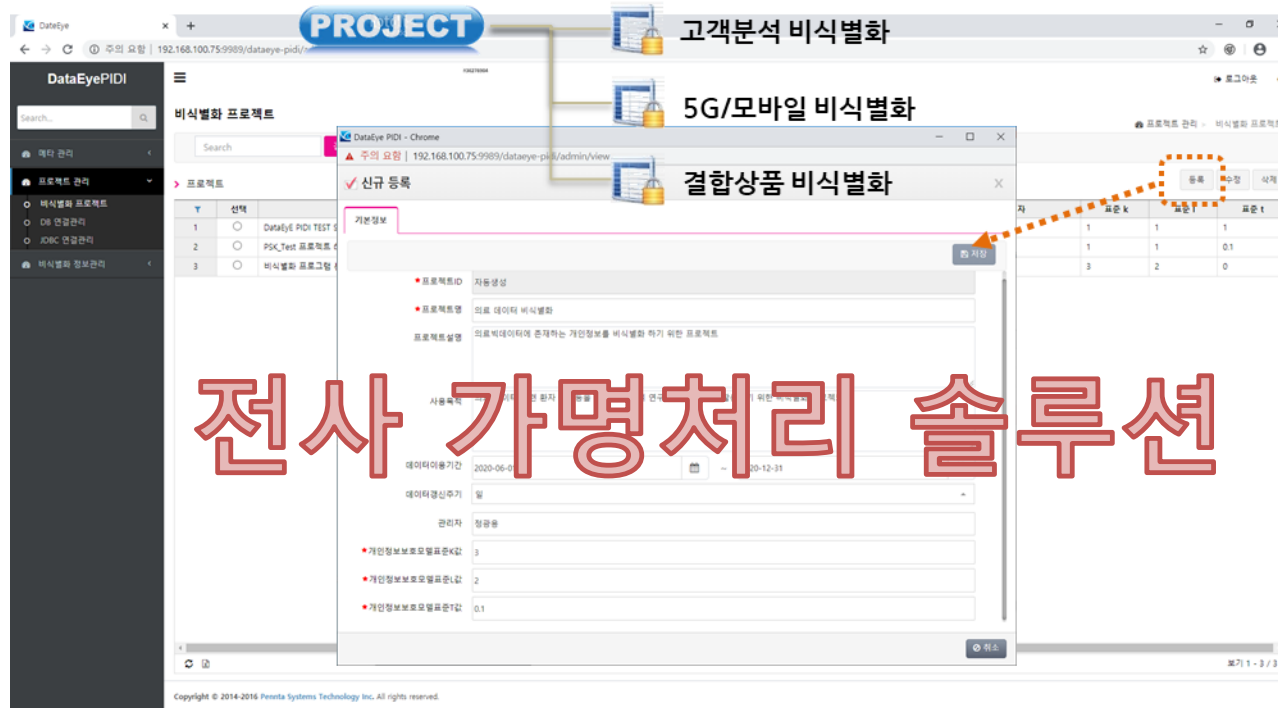


## 대용량 처리 아키텍처



# 5. 특/장점 - 프로젝트별 관리

비식별화 처리를 위한 단위 프로젝트별로 관리가 가능하여 프로젝트 단위의 사용자 및 권한 적용이 가능합니다.



## 주요 기능



- 관리대상의 비식별화 프로젝트 관리정보를 등록, 수정, 삭제할 수 있습니다.
- 개인정보 비식별화 프로젝트는 "프로젝트" 단위로 관리되며 하나의 프로젝트는 다수의 테이블을 포함합니다.
- 프로젝트 관리정보로는 프로젝트명, 사용목적, 데이터 이용기간, 관리자, 표준k, 표준, 표준 정보를 입력하여야 합니다.
- 개인정보 보호 모델(k-익명성, 1-다양성, t-근접성) 값을 지정하여 비식별 조치 적정성 평가를 하게 됩니다.

**감사합니다**