데이터 품질 진단, 이제 버튼 한 번으로

DQube 3.5.1 — 자동 진단 · 도메인 자동매핑 · 보정 진단

㈜엠아이티소프트 | 2026-10

DB를 고르고 [자동진단 시작]만 누르면 프로파일링 · 룰 지정 · 진단 · 보정까지 자동으로 진행합니다
담당자는 대시보드에서 진행 상황과 결과만 확인하면 됩니다

목차

PART 1 새로워진 DQube

01 진단 담당자의 반복 작업을 자동화했습니다
02 핵심 기능 6가지
03 기존 DQube와 무엇이 달라졌나
04 자동 진단, 이렇게 진행합니다
05 기대 효과

PART 2 상세 기능 설명

A 기능 개선 내역
B 자동매핑 실행 절차와 결과
C 자동매핑 대상과 판정 기준
D 진단대상과 자동 진단
E 보정 진단
F 설정 기본값 · 데이터 기준

01

PART 1 · 새로워진 DQube

진단 담당자의 반복 작업을 자동화했습니다

기존 방식 — 사람이 하나씩

- 테이블 · 컬럼마다 도메인과 룰을 직접 지정
- 진단 실행 후 오류 데이터를 하나씩 확인
- 오류가 많으면 룰을 고치고 다시 진단 — 반복
- DB가 많으면 DB마다 같은 작업을 되풀이
- 중간에 끊기면 처음부터 다시 실행

DQube 3.5.1 — 버튼 한 번으로 자동

- DB를 고르고 [자동진단 시작]만 누르면 끝
- 컬럼명과 실제 데이터를 분석해 도메인 · 룰 자동 지정
- 오류율이 높은 컬럼은 룰을 보완해 자동으로 다시 진단
- 여러 DB를 동시에, 멈춘 곳부터 이어서 진행
- 담당자는 대시보드에서 진행 상황과 결과만 확인

진단 준비부터 룰 보완까지 반복 작업을 줄여, 담당자는 결과 검토와 품질 개선에 집중할 수 있습니다

02

PART 1 · 새로워진 DQube

핵심 기능 6가지

1 자동 진단

프로파일링 → 룰 세팅 → 진단 → 보정까지 DB별로 자동 진행. 중지 · 이어서 · 비정상 종료 복구 지원

2 도메인 자동매핑

컬럼명(한글 · 영문) 사전과 실제 데이터를 함께 분석해 도메인과 룰을 지정

3 보정 진단

오류율이 높은 컬럼의 오류 데이터를 분석해 진단 룰을 보완하고 다시 진단.
목표 오류율 0.001% 달성

4 진단 현황 대시보드

메인 화면에서 DB별 오류율 · 진단 수행률 · 품질지표별 결과 · 일자별 추이를 한눈에

5 대규모 · 동시 진단

여러 DB를 동시에 자동 진단, 동시 진단 실행 수 설정, DB 그룹 · 즐겨찾기 · 표시 순서

6 지원 DBMS 12종

Oracle · Tibero · SQL Server · MySQL/MariaDB · PostgreSQL · DB2 · CUBRID · Altibase · Sybase ASE · Sybase IQ · SAP HANA · Firebird

03

PART 1 · 새로워진 DQube

기존 DQube와 무엇이 달라졌나

작업	기존 DQube	DQube 3.5.1
도메인 · 룰 지정	컬럼마다 직접 지정	컬럼명 + 실제 데이터 분석으로 룰셋 자동 지정
룰 오류 보완	오류 확인 후 룰을 직접 수정	보정 진단이 오류 데이터를 분석해 룰 보완 · 재진단
진단 실행	DB · 테이블 단위로 하나씩	여러 DB를 골라 한 번에 자동 진행
진행 확인	화면마다 따로 조회	메인 대시보드 · 자동진단 관리 화면에서 확인
중단 대응	처음부터 다시 실행	멈춘 단계부터 이어서 진행
일괄 등록	양식에 맞춰 직접 입력	업무규칙 · 제외 테이블 불러오기(엑셀 · CSV), 업무규칙 자동생성
백업 · 복구	버전이 다르면 복구 오류 가능	이전 버전 백업도 구조 자동 보정, 실패하면 원래대로
지원 DBMS	7종	12종 (Altibase · Sybase ASE · Sybase IQ · SAP HANA · Firebird 추가)

04

PART 1 · 새로워진 DQube

자동 진단, 이렇게 진행합니다

1 DB 등록

DB 접속 정보를 등록하고 스키마 정보 수집 · 진단 제외 테이블 설정

담당자

2 자동진단 시작

자동진단 관리 화면에서 진단할 DB를 체크하고 [자동진단 시작]

담당자

3 자동 진행

프로파일링 → 도메인 · 룰 지정 → 진단 → 보정 진단 까지 자동

DQube

4 결과 확인

대시보드에서 오류율 · 품질지표별 결과 확인, 오류 데이터 검토

담당자

- 담당자가 할 일은 등록 · 시작 · 확인뿐 — 3단계는 DQube가 알아서 진행합니다
- 여러 DB를 함께 체크하면 동시에 진행하고, 진행률은 DB별로 보여줍니다
- 중지 · 이어서 진행이 가능하고, 프로그램이 종료돼도 다음 실행 때 멈춘 단계부터 이어갑니다
- 오류율이 높은 컬럼만 골라 [컬럼자동진단]으로 다시 진단할 수 있습니다

05

PART 1 · 새로워진 DQube

기대 효과

버튼 한 번

진단 준비부터 보정까지

프로파일링 · 룰 지정 · 진단 · 보정을 자동으로

0.001%

목표 오류율 관리

오류율이 높은 컬럼은 자동으로 보완해 다시 진단

12종

지원 DBMS

공공 · 민간 주요 DBMS를 한 도구로 진단

8분 → 1초

진단 대상 조회

테이블 3,192개 DB 기준, 대상 컬럼 정보 조회 시간

- 진단 기간 단축 : 반복되는 도메인 · 룰 지정과 재진단을 자동화해 담당자 작업 시간을 줄입니다
- 진단 품질 향상 : 컬럼명뿐 아니라 실제 데이터를 근거로 룰을 정하고, 판정 근거를 컬럼마다 기록합니다
- 대규모 진단 대응 : 기관별 수십 개 DB도 그룹 · 즐겨찾기로 관리하고 동시에 진단합니다
- 안전한 이관 : 이전 버전 백업도 복구 시 구조를 자동으로 맞추고, 실패하면 원래 데이터로 되돌립니다

PART 2

상세 기능 설명

자동 진단 · 도메인 자동매핑 · 보정 진단의 처리 기준

A 기능 개선 내역
B 자동매핑 실행 절차와 결과
C 자동매핑 대상과 판정 기준
D 진단대상과 자동 진단
E 보정 진단
F 설정 기본값 · 데이터 기준

A

PART 2 · 기능 개선 내역

① 자동 진단 · 자동진단관리

구분	수정 내용	효과
자동 진단 흐름	① 스키마 수집 확인 → ② 프로파일링 → ③ 룰 세팅 → ④ 진단 실행 → ⑤ 룰 보정(선택)	룰 세팅부터 보정까지 자동
자동진단관리 화면	DB 체크 후 시작·중지·이어서, DB별 상태·단계·진행률·오류율, 정렬·필터, 자동 새로고침	한 화면에서 진행 관리
다중 DB 동시 진행	동시 자동진단 DB 수만큼 준비 단계 동시 진행, 동시진단실행수만큼 병렬 처리	여러 DB 진단 시간 단축
진행률	룰지정 컬럼 + 참조 + 업무규칙 전체 대비, 이어서 진행해도 완료분 포함	전체 진행 상황 파악
컬럼자동진단	컬럼진단 화면에서 고른 컬럼만 ③ 룰 세팅 → ④ 진단 → ⑤ 보정	문제 컬럼만 빠르게 재진단
중지 · 이어서	중지 후 다른 DB 대기 멈춤 해소, 비정상 종료 시 그 단계부터 이어서	멈춤·반복 종료 방지
참조 · 업무규칙	복원 DB 자동 진단에서도 참조(RI)·업무규칙(BR) 진단	진단 누락 방지
Event Log	DB별 탭(즐겨찾기 우선·등록순), 새 로그만 추가, 열면 마지막 줄로 이동	깜빡임 없이 진행 확인
준비 구간 조회	대상 컬럼 정보를 DB 단위로 한 번에 조회, 프로파일링 중복 실행 제거	테이블 3,192개: 약 8분 → 1초

A

PART 2 · 기능 개선 내역

② 도메인 자동매핑 · 룰

구분	수정 내용	효과
판정 순서	컬럼명 도메인 + 90% 맞는 룰이면 확정, 아니면 5,000건 데이터 분석	컬럼명과 데이터를 함께 사용
룰 결정	dq_rule 기본 룰의 적합률(90%)·우선순위로 지정	도메인별 고정 룰 대신 데이터 기준
룰 자동생성	맞는 룰이 없으면 대표 형식 90% 이상일 때 "(자동생성)" 룰 생성	적합률 0% 룰 지정 방지
데이터 타입 룰	판정 불가 번호·금액·수량·율 : 문자형 NO(기호포함) B00069, 숫자형 B00068	엉뚱한 룰로 오류율 100% 방지
코드 자동 등록	값 종류(NDV, 통계정보 우선)가 기준(100) 미만이면 사용자 코드 룰 등록	코드 컬럼도 자동 진단
날짜 형식	초 뒤 소수 자리(.000) 인식, 날짜 오판(OUT_DATE·SAL_YYYYMM) 수정	날짜 컬럼 정확도 향상
자동매핑결과	[컬럼명우선/데이터우선] + 선정사유를 AUTO_NOTE 에 기록, 컬럼진단 탭에 표시	판정 근거 확인
기본 룰 추가	B00061~62 위도·경도, B00063~65 M/F·m/f·t/f, B00066~67 년.월, B00068~69	기본 룰 부족 보완
선택 컬럼 재판정	[도메인자동매핑]으로 고른 컬럼은 룰·사유와 관계없이 다시 판정 (확인창)	개별 컬럼 수정 쉬움

A

PART 2 · 기능 개선 내역

③ 보정 진단

구분	수정 내용	효과
보정 진단	진단 후 오류율 0.001% 초과 컬럼의 룰을 자동 보정하고 재진단	목표 오류율 0.001% 이하
번호	정상 값 + 오류 값 + 지금 룰 범위를 모두 받는 정규식으로 넓힘	이전 룰로 되돌아가지 않음
여부 → 코드	값 20종 이하면 사용자 코드 룰로 전환	여부 이외 값 수용
코드	자동 등록한 사용자 코드에 새 값을 더해 다시 등록 (기준 미만)	새 코드 값 수용
금액 · 수량 · 율	B00068 수량(음수포함)으로 교체	범위 오류 해소
설정 · 실행	환경설정 자동 수행·최대 횟수(3), 관리 화면 [선택 DB 보정 진단]	자동 / 수동 선택
기록 · 안전장치	자동매핑결과·로그에 기록, 기존 룰은 고치지 않음, 보정 룰은 자동매핑 후보 제외	추적 가능, 다른 컬럼 영향 없음

A

PART 2 · 기능 개선 내역

④ 화면 · 대시보드 · DBMS · 백업

구분	수정 내용	효과
메인 대시보드	DQube 메인 화면(WebView2), 로고 클릭으로 이동, 보고 있을 때만 갱신(10초)	진단 현황을 첫 화면에서
오류율 표시	DB별 전체 오류율, 상세 데이터 오류율(소수 4자리), 테이블별 오류율·정렬	오류율 목표 관리
DB별 현황 순서	자동 진단 → 전체 오류율 → 진단대상 테이블 → 컬럼 → 업무규칙 수(참조 포함)	중요 지표 우선
DB 목록 관리	그룹·진단대상·표시 순서·즐겨찾기, 원천 DB 접속 없이 저장	복구 DB도 목록 정리
DBMS 추가	Altibase, Sybase ASE, Sybase IQ, SAP HANA, Firebird + 드라이버 동봉	지원 DBMS 12종
일괄 등록	업무규칙·제외 테이블 [불러오기](엑셀·CSV), 업무규칙 자동생성(개수 지정)	등록 작업 단축
백업 · 복구	구조 차이 자동 보정, DQ_COLUMN·DQ_RULE 선택 시 룰 테이블 3개 함께, 메타 테이블 8개 추가	이전 백업도 안전하게 복구
기준정보 보정	시작 시 누락 기본 룰·룰 정규식·제외사유 보정 (확인 후 진행)	복구 DB 현행화
안정성	진단 중지 · 화면 오류 수정, CUBRID 대용량 진단 오류 대응	안정적인 진단

B

PART 2 · 자동매핑 실행 절차와 결과

자동매핑은 6단계로 컬럼마다 판정하고 룰셋을 정합니다

1 대상 선정

복구·룰셋 지정·제외 사유 컬럼 제외 (고른 컬럼은 강제 재판정)

2 데이터 조회

NULL 제외 최대 5,000건, 5만 건 초과는 블록 표본

3 인코딩 확인

BASE64·HEX 해시면 "암호화 컬럼"으로 끝

4 도메인 판정

컬럼명 도메인 + 맞는 룰 → 확정, 아니면 데이터 분석

5 룰 결정

90% 룰 → 자동생성 → 데이터 타입 룰, 코드는 사용자 코드

6 저장

룰셋·제외사유·자동매핑여부·자동매핑결과 저장

실행 위치

- 자동 진단: ③ 룰 세팅 단계에서 진단대상 테이블 전체에 자동 실행 (동시진단실행수만큼 병렬)
- 화면: 테이블 설정 [도메인자동매핑](테이블 단위), 컬럼진단 [도메인자동매핑] · [컬럼자동진단](고른 컬럼만)
- 판정 근거는 자동매핑결과(AUTO_NOTE)에 남고, 컬럼진단 화면 '자동매핑결과' 탭에서 확인
- 프로파일링(샘플 3개)은 판정 근거로 쓰지 않음 — 룰 세팅 때 원천 DB에서 다시 읽어 판정

B

PART 2 · 자동매핑 실행 절차와 결과

판정 결과에 따라 룰셋 · 자동매핑여부 · 제외사유가 정해집니다

판정 결과	룰셋	자동매핑여부	제외사유	진단대상	다음 자동매핑
90% 이상 맞는 기본 룰 있음	그 룰 (등급·적합률 순)	Y	없음	대상	제외 (룰셋 있음)
맞는 룰 없음 + 대표 형식 90% 이상	"(자동생성)" 룰	Y	없음	대상	제외 (룰셋 있음)
번호·금액·수량·율 판정 불가	문자형 B00069 / 숫자형 B00068	Y	없음	대상	제외 (룰셋 있음)
날짜·여부, 자동생성 불가	가장 잘 맞는 룰 (보정 대상)	Y	없음	대상	제외 (룰셋 있음)
위도 · 경도 (이름 + 숫자)	B00061 / B00062	Y	없음	대상	제외 (룰셋 있음)
코드, 값 종류 기준 미만	사용자 코드 "(자동생성)"	Y	없음	대상	제외 (룰셋 있음)
코드, 값 종류 기준 이상	없음 (사용자 지정)	Y	코드세팅필요	제외	제외 (사유 있음)
데이터 없음	없음	비움	미사용컬럼(데이터없음)	제외	제외 (기본 제외사유)
내용 (자유 텍스트)	없음	비움	비정형 텍스트컬럼(패턴 없음)	제외	제외 (기본 제외사유)
인코딩 (BASE64 · HEX)	없음	비움	암호화 컬럼	제외	제외 (기본 제외사유)
환경설정에서 꺼진 도메인	없음	비움	없음	제외	다시 판정

※ 추천도메인은 스키마 수집 때 값 그대로 둡니다. "(자동매핑)"으로 끝나는 사유는 다음 실행 때 지우고 다시 판정합니다.

C

PART 2 · 자동매핑 대상과 판정 기준

세 가지에 하나라도 해당하면 자동매핑하지 않습니다

1 제외사유가 적힌 컬럼

기본 제외사유 6가지와 사용자가 적은 "검토 필요", "코드세팅필요" 같은 사유 모두. 단, "(자동매핑)"으로 끝나는 사유는 다시 판정

2 복구된 컬럼

restore_yn = Y 인 컬럼. 룰셋·사유·샘플을 바꾸지 않음

3 룰셋이 지정된 컬럼

사용자 등록이든 자동매핑이든 이미 룰셋이 있으면 그대로 둠 (진단대상)

기본 제외사유 6가지
비정형 텍스트컬럼(패턴 없음)
암호화 컬럼
미사용 컬럼 · 미사용컬럼(데이터없음)
범정부 진단기준에 해당안됨
비대상 데이터타입
(기준정보 컬럼 : 룰 지정만 제외)

예외 : 다시 판정하는 경우

- 컬럼진단에서 고른 컬럼 ([도메인자동매핑]·[컬럼자동진단])
- 환경설정 '룰세팅·자동매핑 완료 컬럼과 복구 컬럼도 다시 판정' 체크 시 ②·③도 판정
- 문자형·숫자형만 판정 (날짜형·LOB 제외)

C

PART 2 · 자동매핑 대상과 판정 기준

컬럼 상태별 처리 — 위에서부터 처음 맞는 줄을 따릅니다

컬럼 상태 (위에서부터 차례로 확인)	자동매핑	제외사유	진단대상
복구된 컬럼 (restore_yn = Y)	제외	그대로 유지	룰셋이 있으면 대상
룰셋 있음 (사유 유무 무관)	제외	그대로 유지	대상
기본 제외사유 6가지 + 룰셋 없음	제외	그대로 유지	제외
"검토 필요" 등 사용자 사유 + 룰셋 없음	제외	그대로 유지	제외
사유 없음(또는 "(자동매핑)") → 룰 지정 (적합 · 자동생성 · 데이터 타입)	매핑, 룰셋 지정	비움	대상
사유 없음 → 코드, 값 종류 기준 미만	사용자 코드 등록	비움	대상
사유 없음 → 코드, 값 종류 기준 이상	판정만 함	"코드세팅필요"	제외
사유 없음 → 데이터 없음 · 내용 · 인코딩	판정만 함	기본 제외사유 기록	제외
사유 없음 → 꺼진 도메인으로 판정	판정만 함	없음 (자동매핑여부도 비움)	제외
컬럼진단에서 고른 컬럼	강제 재판정	지우고 새로 판정	판정 결과에 따름

C

PART 2 · 자동매핑 대상과 판정 기준

도메인 판정 순서와 룰 결정 순서

1 **컬럼명 + 룰** — 컬럼명(한글·영문)이 사전(std_domain)에 있고 90% 맞는 룰이 있으면 확정

2 **컬럼명 + 자동생성** — 컬럼명 도메인인데 맞는 룰이 없으면 데이터로 룰 자동생성

3 **여부** — 고유값 4개 이하, 하나의 여부 쌍(Y/N·0/1·M/F 등) 99% 이상

4 **날짜** — 날짜 값 90% 이상 (소수 초 .000 포함)

5 **위도 · 경도** — 이름이 위도·경도이고 숫자 데이터 → 수량

6 **데이터만** — 문자+공백 → 내용, 정수 키 → 번호, 일련번호 → 번호, 서술형 → 내용

7 **코드 / 번호** — 고유값 비율 50% 초과면 번호, 아니면 코드

룰 결정 (위에서부터)	지정 룰
① 90% 이상 맞는 기본 룰	등급 → 적합률 → 범위 → 순서
② 대표 형식 90% 이상	"(자동생성)" 룰
③ 데이터로만 정한 번호·금액·수량·율	문자형 B00069 / 숫자형 B00068
④ 그 밖 (날짜·여부 등)	가장 잘 맞는 룰 → 보정 대상
코드 도메인	값 종류 < 기준 → 사용자 코드
위도 · 경도	B00061 / B00062

※ NO(기호포함) B00069 와 보정 진단이 만든 "(자동보정)" 룰은 적합률 후보에서 제외 (다른 컬럼 판정에 영향 없음)

D

PART 2 · 진단대상과 자동 진단

진단대상은 룰셋 지정 여부 하나로 정합니다

진단대상 테이블
테이블 설정에서 진단대상으로 둔 테이블. 0건 테이블 제외는 선택

+

룰셋이 지정된 컬럼
제외사유와 관계없이 룰셋이 있으면 진단. 진단에서 빼려면 룰셋을 지움

=

진단대상 컬럼
진단 실행 · 대시보드 · 오류율의 기준

대시보드 지표	정의
진단대상 컬럼	룰셋이 지정된 컬럼 (모수에서 기본 제외사유 6가지 제외)
전체 오류율	오류 건수 ÷ 진단 건수 (룰지정 컬럼의 최근 진단 결과), 기준 0.001% 이하
진단 수행률	진단 완료 ÷ 진단대상 컬럼
업무규칙 수	등록된 업무규칙 + 참조 규칙
진단대상 제외	룰셋이 없는 컬럼 (제외사유별로 집계)

D

PART 2 · 진단대상과 자동 진단

자동 진단은 5단계를 DB별로 자동 진행합니다

1 스키마 확인

수집된 진단대상 테이블이 있는지 확인 (수집·제외 테이블은 담당자가 먼저)

2 프로파일링

데이터 건수와 샘플 3개 수집 (병렬)

3 룰 세팅

도메인 자동매핑 (0건 테이블 제외는 선택)

4 진단 실행

룰지정 컬럼 + 참조 + 업무규칙 진단

5 룰 보정

환경설정에서 켠 경우, 오류율 0.001% 초과 시

- 자동진단관리 화면에서 DB를 체크하고 시작 — 테이블 설정 [자동진단 시작]도 이 화면을 엶 (DB 선택 불필요)
- 동시 자동진단 DB 수만큼 준비 단계를 동시에, 동시진단실행수만큼 프로파일링 · 룰 세팅 · 진단을 병렬 처리
- 진행률: 룰지정 컬럼 + 참조 + 업무규칙 전체 대비 (이어서 진행해도 완료분 포함)
- 컬럼진단 [컬럼자동진단]: 고른 컬럼만 ③ → ④ → ⑤ (대상 컬럼은 std_auto_job_col 에 저장)
- 진행 상황: 자동진단관리 그리드 · Event Log(DB별 탭) · 메인 대시보드 / 이력 std_auto_job · std_auto_job_log (90일)

PART 2 · 진단대상과 자동 진단

D

중지 · 다시 시작 · 비정상 종료 처리

사용자 [중지]

- 진행 중인 단계에 중지 요청, 대기 중인 진단은 취소
- 원천 DB 조회가 끝나야 멈춤 — 30초가 지나면 백그라운드에서 마무리하고 다음 DB 시작
- 끊긴 진단은 NG로 저장하지 않음 (이전 결과 유지)

다시 시작

- [이어서]: 멈춘 단계부터 진행
- 진단 실행 단계는 이미 진단한 컬럼을 건너뛰고 남은 컬럼만
- 컬럼자동진단 작업도 고른 컬럼 그대로 이어서

프로그램 종료

- 정상 종료: "프로그램 종료로 중지" 기록, 다음 실행 때 그 단계부터 자동 진행
- 비정상 종료: 이어서 진행하되 같은 단계에서 3회 반복되면 오류로 중단

E

PART 2 · 보정 진단

전체 오류율 0.001% 이하를 목표로 룰을 보정하고 다시 진단합니다

1 진단 완료

④ 진단 실행이 끝남

2 오류율 확인

전체 오류율 0.001% 이하면 끝

3 대상 선정

컬럼 오류율 0.001% 초과 컬럼

4 룰 보정

오류 값을 받는 새 룰로 교체

5 보정 재진단

룰을 바꾼 컬럼만 다시 진단

도메인	보정 방식	보정하지 않고 검토 대상으로 남김
번호	정상 값 5,000건 + 오류 값 + 지금 룰 범위를 모두 받는 정규식 (넓히기만)	정규식 미지원 DBMS, 넣을 수 없는 문자, 더 넓힐 수 없음
여부	값 20종 이하면 사용자 코드 룰로 전환	값 20종 초과
코드 (자동 등록 코드)	기존 코드 + 새 값으로 사용자 코드 다시 등록	값 종류 기준 이상, 따옴표·줄바꿈 값
금액 · 수량 · 율	B00068 수량(음수포함)으로 교체	숫자가 아닌 값
날짜 · 내용 · 사용자 코드 · 이메일	보정하지 않음	—

※ 3 ~ 5 를 전체 오류율이 기준 이하가 되거나 최대 반복 횟수(기본 3회)까지 반복합니다.

E

PART 2 · 보정 진단

실행 방법 · 기록 · 안전장치

자동 수행

환경설정 > 일반설정 > 자동진단 탭 "자동 진단 후 보정 진단 자동 수행" (기본 사용 안 함), 최대 반복 1~10회 (기본 3회)

수동 실행

자동진단관리 화면에서 DB 체크 → [선택 DB 보정 진단]. 진단 결과가 있는 DB만

컬럼자동진단

컬럼진단 [컬럼자동진단]은 고른 컬럼만 오류율 계산 · 보정 · 재진단

기록

자동매핑결과에 "[룰 보정 n차] 이전 룰 → 새 룰 / 오류율" 기록, 재진단 후 오류율 채움. 로그 단계 CORRECT · RECHK

룰 관리

기존 룰은 고치지 않고 새 룰 "(자동보정)"을 만들어 컬럼 룰만 바꿈. 같은 내용 룰이 있으면 재사용

안전장치

번호 정규식은 넓히기만, 보정 룰은 자동매핑 후보 제외 → 재판정 · 반복 시 이전 룰로 되돌아가지 않음

F

PART 2 · 설정 기본값 · 데이터 기준

설정 기본값과 데이터 기준

항목	알아야 할 내용
도메인 자동매핑 대상	신규 설치 기본값은 모든 도메인 켜짐. 끈 도메인은 판정만 하고 자동매핑여부 · 제외사유를 남기지 않음
다시 판정 체크박스	"룰세팅·자동매핑 완료 컬럼과 복구 컬럼도 다시 판정" 체크 시 그 컬럼도 기존 룰을 비우고 다시 판정
코드 도메인 기준	값 종류(NDV) 기준 기본 100 — 미만이면 사용자 코드 자동 등록, 이상이면 코드세팅필요
보정 진단	기본 사용 안 함, 최대 3회. 목표 오류율 0.001% (전체 합계 기준)
동시진단실행수	진단모니터링 · 자동진단관리 같은 값 사용, 진단 · 프로파일링 · 룰 세팅에 모두 적용
데이터 건수 · 값 종류	통계정보를 먼저 쓰고 없거나 0 이면 COUNT(*) / 판정 데이터 5,000건
판정 데이터 범위	5만 건 이하 테이블 · 표본 미지원 DBMS는 값이 있는 앞쪽 5,000건으로 판정 — 드문 코드 값은 빠질 수 있음
코드 컬럼 찾기	제외사유 "코드세팅필요"로 조회해 코드 테이블 룰을 지정

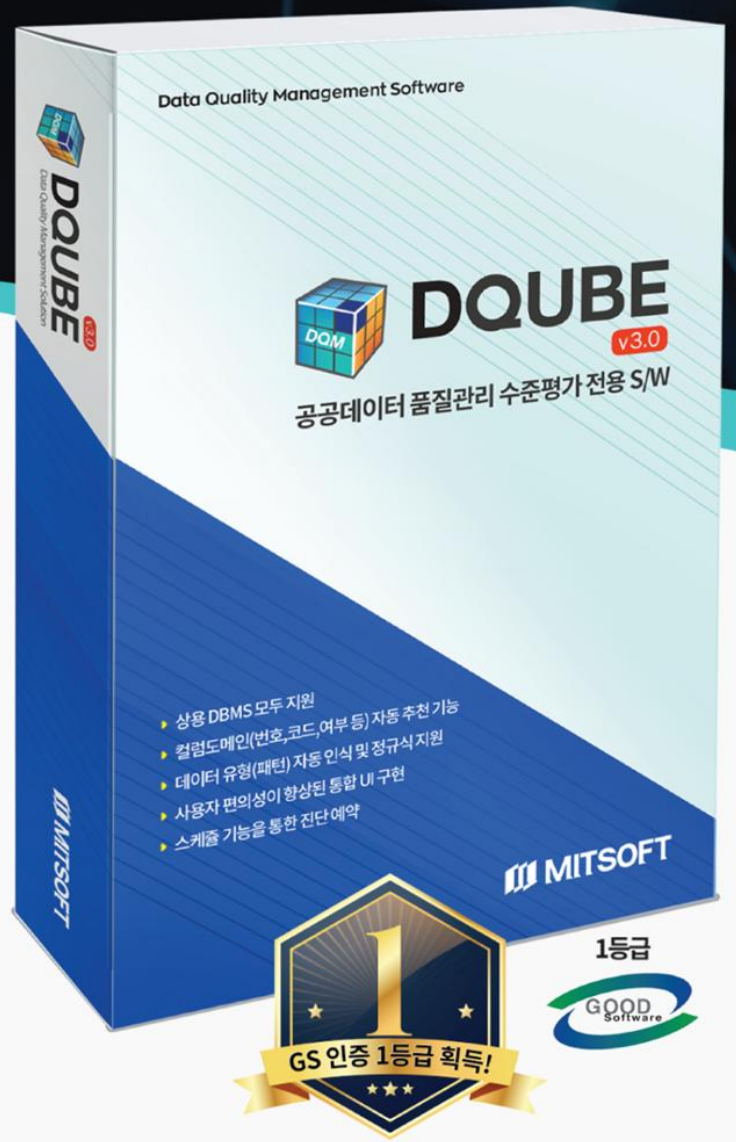
Data Quality Management Software

DQUBE
v 3.0

공공데이터 품질관리 전용 S/W

- 모든 종류의 상용 DBMS 품질 진단 지원
- 프로파일링 및 업무규칙(BR) 진단
- 진단 대상 테이블 및 컬럼 정보 자동 수집 기능
- 컬럼 도메인 (날짜, 번호, 코드,여부 등) 자동 추천 기능 제공
- 진단규칙, 정규식 작성 및 테스트 기능
- 엑셀과 동일한 방식의 편리한 사용자 인터페이스 제공
- 테이블 데이터 검색 기능 내장 (Query 기능을 위한 별도의 S/W불필요)
- 진단 결과 및 이력 정보 한 화면에서 확인 가능
- 오류 데이터 한 화면에서 즉시 확인 가능
- 스케줄링 기능을 통한 예약 진단
- 진단 대상 데이터 Sampling 기능 제공
- 진단결과보고서 출력 및 저장 기능 (오류데이터 포함여부 선택 가능)
- 진단결과보고서 "수준평가" 증적 자료로 제출 가능

품질진단결과보고서는 수준평가 증적 자료로 활용



㈜엠아이티소프트 www.mitsoft.kr 제품 문의 : sales@mitsoft.kr (Tel : 02-719-0380, Mobile : 010-2642-0380)