

Data Integrity & Governance

글로벌 DI 규제 대응 전략 세미나

- Data Integrity
- Regulation
- Automation
- AI



📅 2026년 5월 26일 (화) 09:30 - 17:30

📍 서울대학교 시흥캠퍼스 교육협력동 109호

METTLER TOLEDO
Your Partner in Precision™

경기도 서울특별시 SNU
제약바이오인력양성센터
경기시흥글로벌바이오캠퍼스

Data Integrity to Data Value

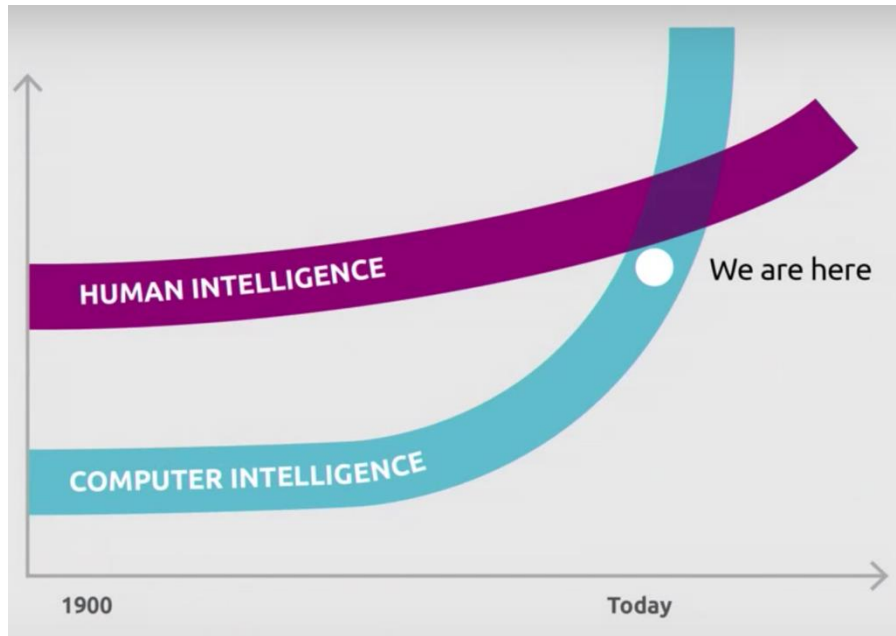
AI 시대, 품질관리 경쟁력은 데이터 활용 역량에서 결정된다

김화종

한국제약바이오협회 K-MELLODDY 사업단장

목차

- 현장의 문제
- AI는 무엇을 할 수 있나
- 규제 친화적 AI
- Data Integrity에서 Data Value로



현장의 문제

01

반복적·시간소모·데이터 기반 → AI로 해결 가능

현장의 문제

문서 작성 부담

Deviation 1건에 보고서·CAPA·SOP 개정까지 수십 장의 문서를 사람이 직접 작성

OOS·이탈 사후 대응

이미 발생한 OOS를 뒤늦게 발견 — 원인 추적과 재발 방지에 막대한 시간 소모

데이터는 쌓이는데

LIMS·DCS·MES에 데이터는 넘치지만 실제 의사결정에 연결되지 못하는 정보의 낭비

규제 대응 과부하

FDA·EMA 실사 준비에 전문 인력이 수개월 투입 — 반복 작업에 혁신 역량 낭비

이러한 문제의 공통점: 반복적이고, 데이터가 있고, AI가 해결할 수 있다

데이터는 규제 대상인가, 경쟁 자산인가?

Compliance Era

- 데이터는 규제 의무 · 감사용 증빙
- 수기 기록 관리
- 사후 대응 (Reactive)
- 비용 발생 요인으로 인식
- 부서간/기기간 데이터 분절



AX Era

- 데이터는 전략 자산 · 인사이트 원천
- AI로 자동 분석·실시간 처리
- 사전 예측 (Proactive)
- 비즈니스 경쟁력 핵심
- 효율적인 통합 데이터 관리

AI는 무엇을 할 수 있나

02

예측·추천·최적화·이상탐지·생성

AI의 진화

머신러닝	딥러닝	Foundation Model	Science Foundation Model	Agentic AI	AX
선형모델 트리모델 부스팅 모델	신경망 (CNN,RNN, Transformer, Graph Net)	LLM (OpenAI, Claude, Gemini 사내 LLM)	BioGPT ChemBERT AlphaFold Self-driving Lab	프롬프트 RAG Fine-tuning	전 산업 영역 혁신
예측, 추천 최적제어 예지 보수	신약후보 발굴 리드 최적화 진단, 치료	문서 작성 질의 응답 문서 번역	화합물 예측 단백질 설계 자율실험	규제 대응 업무 자동화 AI 연구자	신약개발, 규제, 임상, 제조 전 과정 변화

데이터(x) → 모델(AI) → 가치(y)

수식, 논리로 설명하기 어려운 복잡한 실세계 문제를, 데이터로부터 학습한 모델로 해결한다

예측

Prediction

$X \rightarrow \text{Model} \rightarrow y$

제조: OOS·배치 품질 예측
신약: 독성·ADMET·PK 예측

추천

Recommendation

$X \rightarrow \text{Model} \rightarrow \text{candidate}$

제조: CAPA 후보, 원인분석 추천
임상: 시험 프로토콜 조건 추천

최적화

Optimization

$X \rightarrow \text{Model} \rightarrow \text{control}$

제조: CPP 조정, 수율·순도 최적화
임상: 용량·투여일정 최적화

이상탐지

Anomaly Detection

$X \rightarrow \text{Model} \rightarrow y$

제조: audit trail 이상, 공정 drift
임상: 부작용 신호 탐지 (PV)

생성

Generation

$X \rightarrow \text{Model} \rightarrow x'$

GMP: SOP·Deviation 보고서
규제: CTD·IB 문서 생성, 검토

멀티모달 AI



서로 다른 형식의 데이터를 공통 잠재 벡터 공간에서 학습하여 파운데이션 모델 개발

규제 친화적 AI

03

규제기관 관점의 검증 가능한 AI

규제 친화적(Trustworthy) AI

일반적 AI

✗ 최고 성능 모델 추구

✗ 블랙박스 모델 한계

✗ 데이터 출처·버전 관리 미흡

✗ 감사 대응 한계



규제 친화적 AI

✓ 불확실성 정량화 (예: confidence score, prediction interval)

✓ 추적가능성 (Data + Model lineage)

✓ 재현성·검증 가능성 내장

✓ 운영 중 성능 저하 모니터링

Human Oversight 원칙: AI는 판단을 보조하고, 최종 품질 판단과 승인은 책임 있는 GMP 담당자가 수행한다

규제기관의 핵심 질문: 이 AI의 COU(Context of Use)는 무엇인가? 어떤 데이터로 개발·검증했는가? 위험도는 어느 수준인가? 운영 중 성능 저하와 변경관리는 어떻게 수행하는가?

Good AI Practice 원칙

GLP	GCP	GMP	GVP
리스크 동물실험 데이터 조작	리스크 임상시험 윤리·데이터 신뢰성	리스크 제조 품질 불균일·오염	리스크 시판 후 부작용 관리 부실
✓ 해결책 실험실 운영 표준화	✓ 해결책 인체 시험 보호·문서화	✓ 해결책 공정·시설·품질 표준화	✓ 해결책 약물감시 체계화

Good AI Practice 원칙

AI에 요구되는 신뢰성 원칙 제시

- AI 모델은 배포 후에도 성능이 변한다 (Data Drift)
- 일부 AI 모델은 의사결정 과정이 불투명할 수 있어, 기존 밸리데이션에 더해 설명가능성, 불확실성, 성능 모니터링이 필요하다.
- AI는 학습 데이터의 편향을 그대로 학습: 잘못된 데이터 → 잘못된 예측 → 환자 피해

2026.1 “Guiding Principles of Good AI Practice in Drug Development” – FDA, EMA 등

10대 원칙: human-centric design, risk-based approach, clear context of use, data governance and documentation, risk-based performance assessment, lifecycle management 등

Data Integrity에서 Data Value로

04

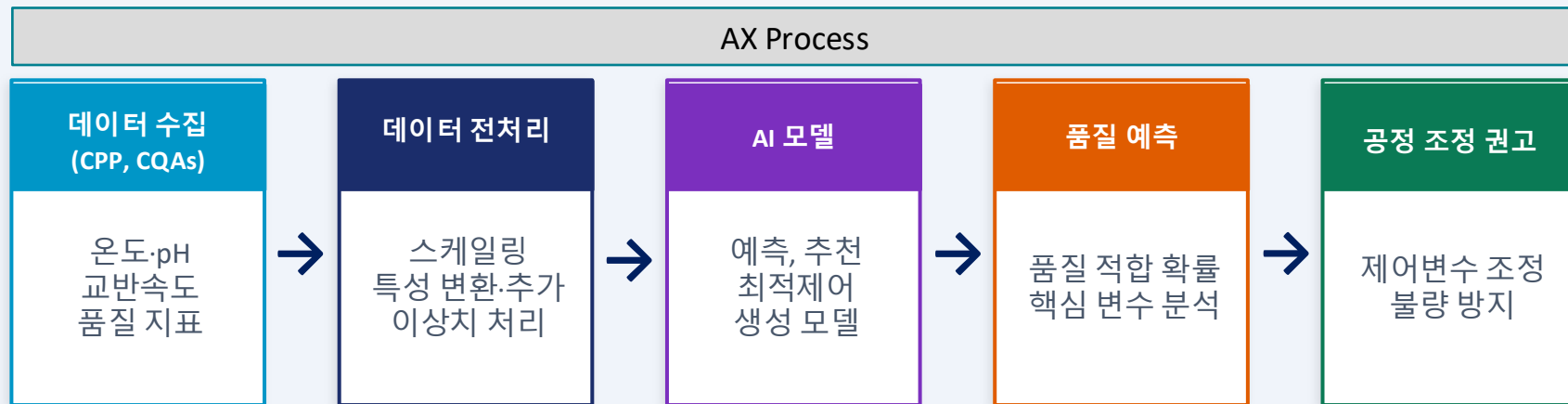
GMP의 AI 전환 전략

Data Integrity의 AX

기존 Data Integrity	➔	AX 도입 이후
수작업 입력·전사 → 오타자·누락 빈발	디지털화	LIMS·DCS 자동 수집 → 전사 오류 감소
OOS 발생 후 사후 발견	이상 탐지	AI 실시간 모니터링 → 이탈 전 선제 경보
담당자가 보고서·CAPA 직접 작성 (수일)	Deviation 처리	LLM 초안 자동 생성 → QA 검토 수행
Audit trail 수동 검토 — 실사 전 집중	감사 추적	AI 상시 분석 → 이상 패턴 자동 플래그
SOP·배치기록 분산 보관 → 버전 혼선	문서 처리	RAG 기반 최신본 실시간 검색·링크 제공
실사 준비에 수개월 전담 인력 투입	규제 대응	실사 대응 패키지, Deviation/CAPA summary, Batch Review Report 준비 지원

QbD (Quality by Design)는 품질을 설계하는 철학

AI는 CPP (공정 파라미터)와 CQA (품질 속성) 관계를 정밀하게 모델링하여 QbD 실행을 지원하고, 위험 기반 공정 조정 의사결정을 보조한다



GMP의 AI 도입 영역

GMP 전 영역에 AI를 접목해 품질관리를 사후 대응(Reactive)에서 사전 예측·최적화(Proactive)로 전환

① 공정 제어 & 품질 예측

- CPP↔CQA 관계를 AI 모델로 정밀 학습
- 배치 초기 데이터로 최종 품질 예측
- 수율 최적화를 위한 공정 제어 지원

② OOS·이탈 조기 감지

- 공정 파라미터 실시간 AI 모니터링
- 이탈 가능성 조기 경고 → 사전 검토
- Audit Trail 이상 패턴 자동 점검

③ 문서화 & CAPA 자동화

- Deviation 보고서 LLM 초안 생성
- SOP·규정 RAG 기반 실시간 검색·링크
- CAPA 후보안 제안 → QA 검토

④ 설비 예지보전

- IoT 센서 + ML로 고장 징후 사전 감지
- Remaining Useful Life(RUL) 예측
- 비계획 다운타임 감소

⑤ 규제 대응 지원

- 실사 대응 패키지 준비
- Batch Review Report 요약
- 21 CFR Part 11 관련 데이터·접근·감사추적 점검 지원

⑥ 공급망 & 원료 품질

- NIR 스펙트럼으로 원료 배치 품질 예측
- 같은 COA라도 공정 성능 차이 식별
- Real-Time Release Testing(RTRT) 지원

Data Integrity는 AI 도입의 전제조건 — 나쁜 데이터로는 우수 AI를 만들 수 없다

사례 ① | 배치 수율 예측 AI (Recordati + Aizon, 2025)

Recordati (Cork, Ireland) + Aizon

ISPE Pharma 4.0™ Conference
2025

● 도전 과제

- 제조 공정의 변동성 요인을 실시간으로 파악하기 어려움
- 배치별 수율 편차가 크고 원인 추적에 시간 소모
- COGS(매출원가) 통제 어려움 → 원가 경쟁력 약화
- 데이터는 존재하지만 의사결정에 연결되지 않는 구조

● AI 솔루션

- AI 데이터 분석 플랫폼(Aizon)으로 공정 변동 요인 파악
- 수율에 영향을 미치는 핵심 공정 파라미터(CPP) 식별
- GxP/21 CFR Part 11 요구사항을 고려한 데이터 기반 공정 분석

공정 데이터를 "규제 증빙"에서 "수율 개선 자산"으로

성과

▲ 1.5%
수율 향상

▼ 2%
매출원가 절감

3개월
파일럿 적용 기간

"데이터를 수율 개선에 활용하면 3개월
파일럿 과제로도 재무 성과 향상"

사례 ② | 제조 데이터 통합과 이상탐지: Eli Lilly의 GMDF 사례

Eli Lilly

GMDF (Global Manufacturing Data Fabric)

● 도전 과제

- LIMS·MES·DCS 등 시스템별 데이터 분절 → 통합 공정 이해 어려움
- 배치 이탈을 QC 완료 후 사후적으로 발견
- 기준 배치 또는 우수 배치 패턴 비교 체계 부재
- 데이터는 있으나 공정 지능(Process Intelligence)으로 연결 안 됨

● AI 솔루션 (Solution)

- **GMDF** 구축: 전 공정 데이터를 단일 통합 데이터 레이어로 연결
- 이상탐지 및 공정 최적화에 활용
- Batch Release by Exception 지원: 기준 충족 배치는 신속 검토하고, 이탈 가능 배치에 QA 검토 역량 집중

Data Integrity가 탄탄해야 기준 배치 비교와 이상탐지가 가능하다

성과

BEFORE

QC 완료 후에야 배치 이탈 발견

AFTER

공정 중 이상 징후 조기 감지 및 경보 지원

데이터 사일로 → 통합 지능
차별화된 제조 데이터 인프라 구축

"배치의 최종 품질은 이미 공정 초기 데이터에 담겨 있다. AI는 그 신호를 더 빠르게 포착하도록 돕는다."

사례 ③ | 설비 예지보전 : Pfizer 사례 등

Pfizer + GE Vernova

글로벌 제약사의 PdM 적용

● 도전 과제

- 설비 고장을 사후 발견 → GMP 배치 무효 + 수익 원 손실
- 정기적 예방보전이 실제 설비 상태와 무관하게 운영
- 비계획 다운타임이 GMP 이탈 기록·규제 대응으로 이어지는 악순환

● AI 솔루션 (Solution)

- Pfizer + GE Vernova: 제조 현장 데이터와 predictive maintenance 를 활용해 장비 상태를 모니터링하고 운영 성능 개선
→ 예방보전 → 예지보전(PdM) 전환:
- 글로벌 제약사 비계획 다운타임 감소 효과 보고
※ 업계 PdM 사례에서 25~35% 수준의 감소 효과가 보고됨

성과

다운타임 감소

생산성·신뢰성 향상

생산성 향상

가동시간 및 생산 안정성 향상

OEE 향상

설비 종합효율 개선

"사후 수리"에서 "예측 기반 계획 정비"는
GMP 컴플라이언스의 중요한 기반이 된다

실패 패턴

AI 도입이 새로운 업무 부담이라고 인식



AI는 문제 해결의 핵심 도구로서 업무를 줄이는 수단으로 재정의

해결할 문제(요구사항)를 명확하게 정의하지 못함



모든 관계자가 참석하여 문제 정의 워크숍 진행

보유한 데이터의 특성과 가치를 파악하지 못함



데이터의 가치는 작은 AI 모델·파일럿을 통해 구체화됨

부서별 독립적 추진



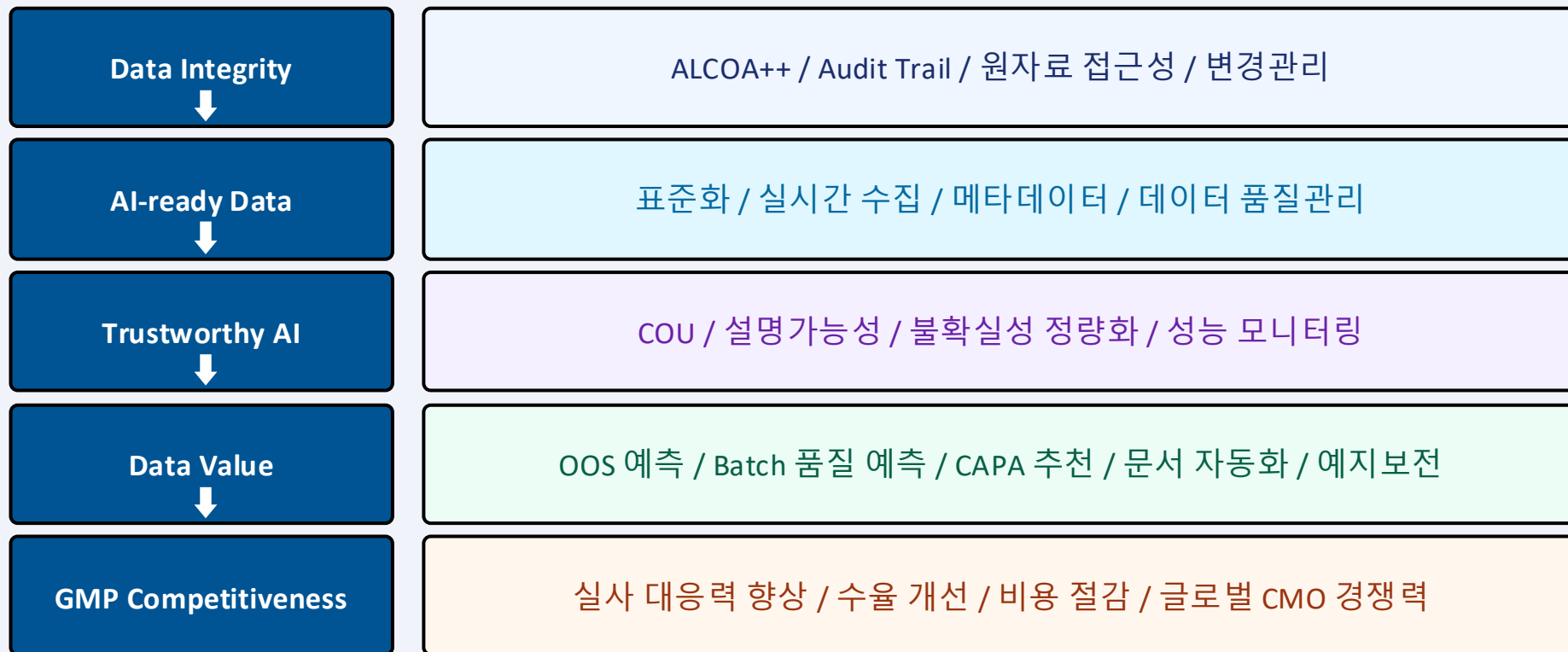
QA/QC/생산/IT 전 부서가 문제 정의와 데이터 기반 협력을 해야 함

AI 추진 역량(기획·운영·검증)을 내재화하지 못함



사내 AI 기획·운영·검증 역량 내재화 필요

Data Integrity to Data Value



결론

AI 시대의 GMP 경쟁력은 데이터를 단순히 보존하는 Data Integrity에서, 데이터를 예측·판단·개선에 활용하는 Data Value로 전환하는 데 있다

AI는 GMP를 실시간·예측형·증거기반으로 만든다

Data Integrity는 AX의 출발점이다. 나쁜 데이터로는 신뢰 가능한 AI를 만들 수 없다

규제기관은 AI 도입에서 COU·검증·설명가능성·생애주기 관리를 요구한다

LLM은 의사결정자가 아니라, 문서화·검색·검토를 돕는 GMP co-pilot

작은 pilot에서 시작해 AI-ready QMS로 단계적으로 전환해야 한다

Q & A

감사합니다