



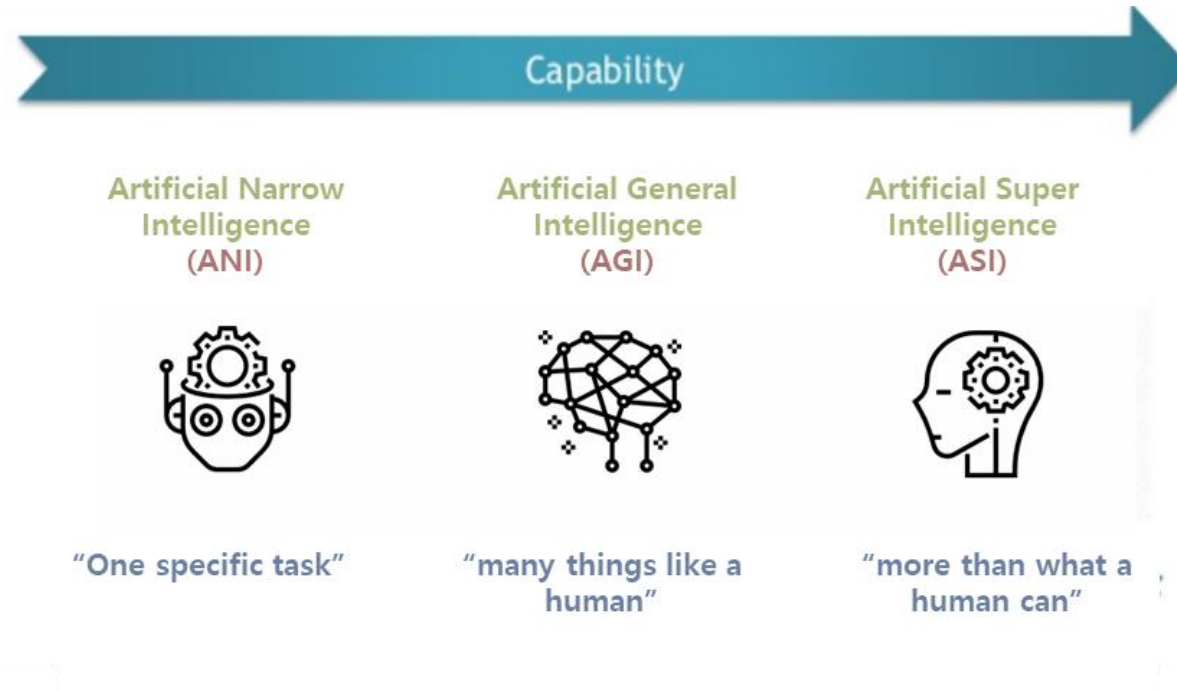
# Data Scientist : 제조 데이터

2025.

- I. 제조 및 품질와 AX 개요
- II. 품질4.0과 데이터 사이언스
- III. 관련 사례

# I. 제조 및 품질과 AX 개요

- AI 개요



- '약인공지능(ANI)', '범용인공지능(AGI)', '초인공지능(ASI)'
- ANI: 주어진 데이터 내에서 주어진 문제를 해결하는 AI
- AGI로의 진입 가속화 중

# I. 제조 및 품질과 AX 개요

- Scaling Laws (Open AI, 2020)

- 컴퓨팅 리소스, 데이터, 모형 크기를 늘릴 수록 성능 개선
- 새로운 능력이 생겨남 (Emergent Abilities)

Scaling Laws for Neural Language Models			
Jared Kaplan *		Sam McCandlish*	
Johns Hopkins University, OpenAI		OpenAI	
jaredk@jhu.edu		sam@openai.com	
Tom Henighan	Tom B. Brown	Benjamin Chess	Rewon Child
OpenAI	OpenAI	OpenAI	OpenAI
henighan@openai.com	tom@openai.com	bchess@openai.com	rewon@openai.com
Scott Gray	Alec Radford	Jeffrey Wu	Dario Amodei
OpenAI	OpenAI	OpenAI	OpenAI
scott@openai.com	alec@openai.com	jeffwu@openai.com	damodei@openai.com

## Abstract

We study empirical scaling laws for language model performance on the cross-entropy loss. The loss scales as a power-law with model size, dataset size, and the amount of compute used for training, with some trends spanning more than seven orders of magnitude. Other architectural details such as network width or depth have minimal effects within a wide range. Simple equations govern the dependence of overfitting on model/dataset size and the dependence of training speed on model size. These relationships allow us to determine the optimal allocation of a fixed compute budget. Larger models are significantly more sample-efficient, such that optimally compute-efficient training involves training very large models on a relatively modest amount of data and stopping significantly before convergence.

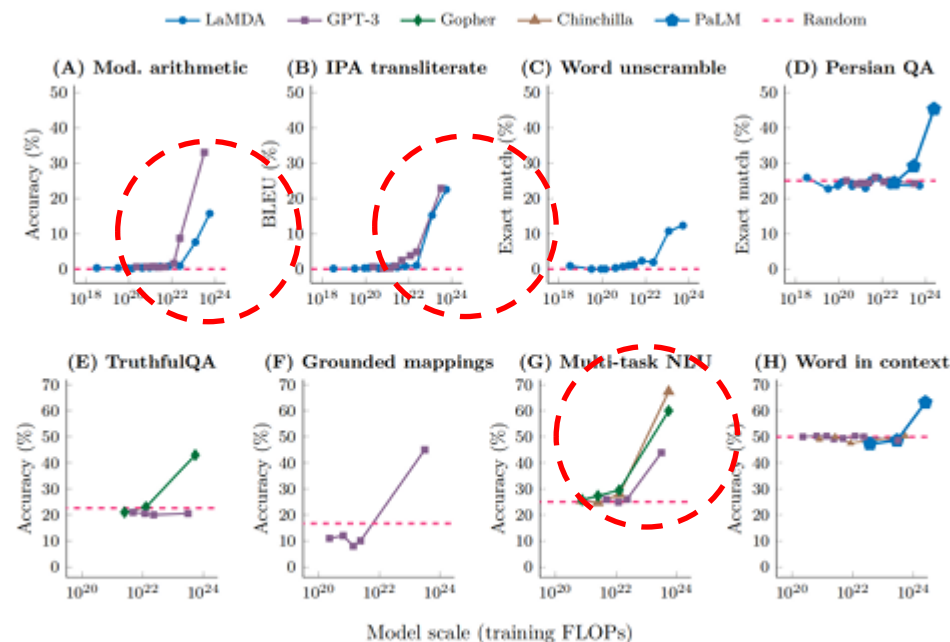


Figure 1: Emergent abilities of large language models. Model families display *sharp* and *unpredictable* increases in performance at specific tasks as scale increases. Source: Fig. 2 from [33].

# I. 제조 및 품질와 AX 개요

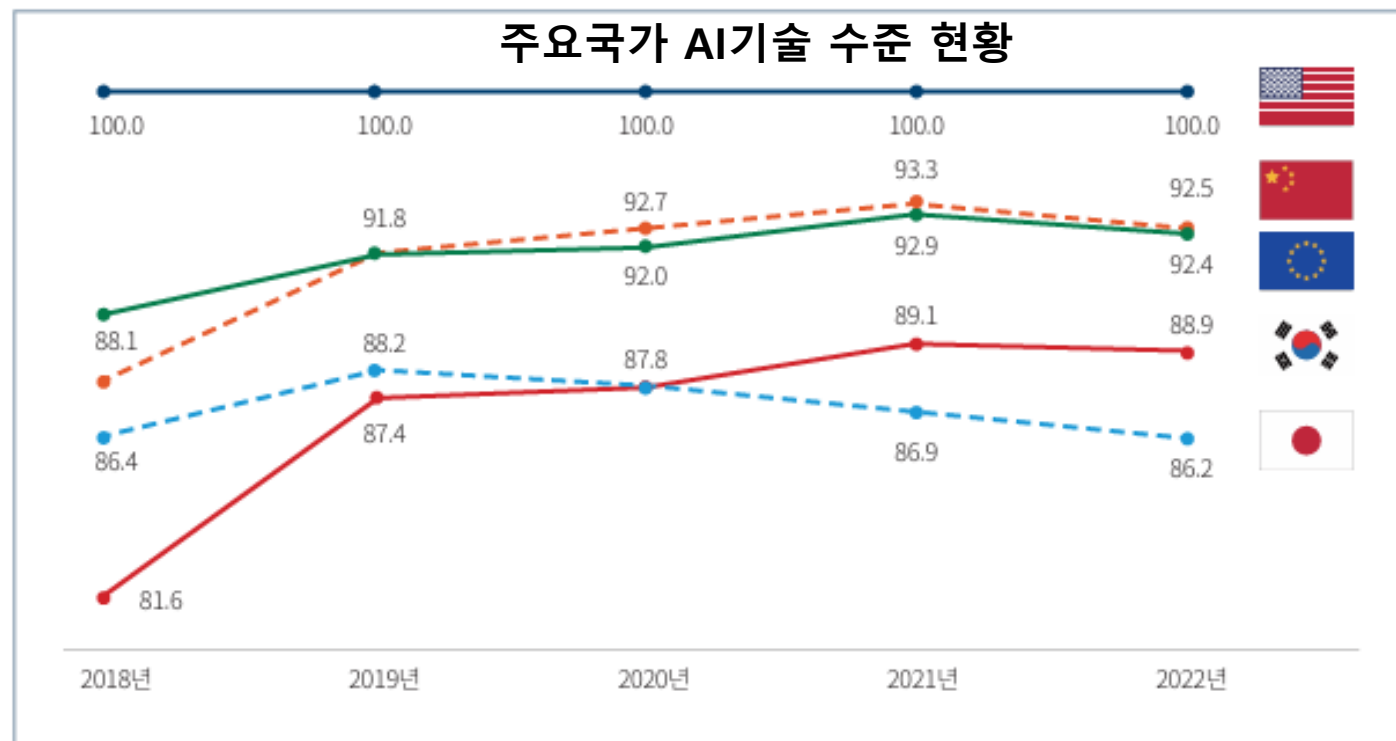
## • AI 역량 비교

- 주요 국가 중 기술 수준이 빠르게 높아지고 있음
- 주요 국가 대비 여전히 기술 격차 존재

(단위: %)

국가	2018년(A)	2019년	2020년	2021년	2022년(B)	증감치(B-A)
한국	80.5	86.2	87.2	87.7	88.0	7.5
미국	100.0	100.0	100.0	100.0	100.0	0.0
일본	86.4	87.2	87.2	85.6	85.7	-0.7
중국	86.8	90.8	92.0	92.0	92.0	5.2
유럽	92.0	93.3	94.3	95.7	95.7	3.7

출처: 정보통신기획평가원, 「ICT 기술수준조사 및 기술경쟁력분석 보고서」 및 「ICT 기술수준조사 보고서」, 각년도 연구자 재구성.



출처: 정보통신기획평가원, 「ICT 기술수준조사 및 기술경쟁력분석 보고서」 및 「ICT 기술수준조사 보고서」, 각년도 연구자 재구성.

Source: 우리나라 및 주요국 인공지능(AI) 기술수준의 최근 변화 추이 (SPRI,2023)



# I. 제조 및 품질와 AX 개요



Source: <https://www.hyundai.co.kr/story/CONT00000000000003259>

## 스마트 팩토리

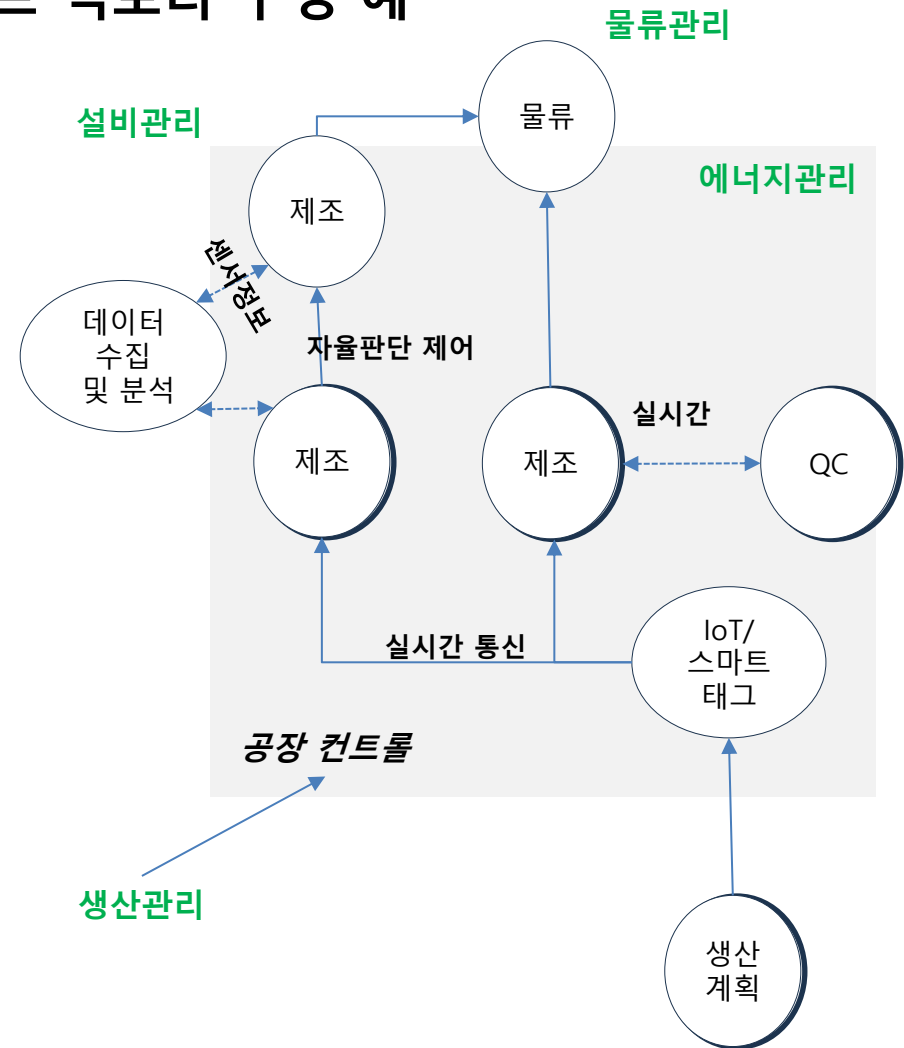
- 생산 현장+디지털 기술: 생산 방식의 혁신과 최적화
- 공장 내 설비와 기계의 센서 설치 및 실시간 데이터 수집
- 데이터 분석을 통한 모니터링, 최적의 생산 방식 도출
- AGV(Automated Guided Vehicle, 무인운반차), 드론, 자율 주행 로봇 등
- 작업 효율성 제고 및 불량률 감소

# I. 제조 및 품질과 AX 개요

## 스마트 팩토리

- 빅데이터 및 AI: 다양한 데이터 활용, AI기반 예지형 서비스
- 디지털 트윈: 가상의 자원과 프로세스 기반(예측)
- Adaptive IoT: IoT 연계로 다양한 센서 등 적용

## 스마트 팩토리 구성 예





# I. 제조 및 품질과 AX 개요

## 스마트 팩토리

ChosunBiz

### [푸드테크 새싹]⑤ “일손 부족한 농촌 ‘불량 과일’ AI가 대신 선별”...에이오팜

곽호재 에이오팜 대표 인터뷰  
군 복무 중 농촌 일손 부족 체감, 전역 후 창업  
창업 2년차, 16억원 투자 유치  
AI 선별 시스템으로 수작업 농산물 대체  
시간당 3만2700개 농산물 처리

이민아 기자  
입력 2022.11.22 07:36

푸드테크 스타트업 에이오팜은 농산물산지유통센터(APC)에서 사람이 일일이 눈으로 보고 골라냈던 불량 농산물을 자동으로 골라내는 인공지능(AI) 선별 시스템을 개발한 창업 2년차 회사다.

곽호재(28) 에이오팜 대표가 군복무를 하며 농촌의 일손 부족을 체감하고, 이를 해결할 아이디어를 구상했다. 그는 이때 만난 박수빈 에이오팜 최고기술책임자(CTO)와 전역 후 지난해 3월 에이오팜 법인을 설립했다.

농촌에는 살아본 적도 없는 인천 출신 곽 대표는 농업이야말로 신기술이 지금 당장 필요한 분야라고 생각하고 있다. 농촌은 신종 코로나 바이러스 감염증(코로나19)으로 외국인 근로자들의 수가 줄면서 심각한 인력난을 겪고 있는데, AI로 부족한 일손을 해결할 수 있다고 믿기 때문이다.

농산물 품질 선별 작업을 인력에 의존할 경우 숙련된 인력도 1시간에 1000개 이상 선별하기 어렵다. 하지만 에이오팜이 개발한 AI 선별 시스템은 시간당 3만2700개의 농산물을 처리한다.

이데일리

2022 카타르 월드컵

기업 전자 자동차 생활 산업 중소기업 아웃도어·캠핑 기업일반

### [테크 Talk]"인공지능 도입하자"...AI 꽃힌 삼성·LG·SK

삼성전자, 웨이퍼 운반 로봇 도입하고 디지털 트윈 적용  
SK하이닉스도 AI로 반도체 계측...불량 감소, 생산성 증가  
LG전자, 자동차부품 성능 검증 AI 플랫폼으로 품질 개선

등록 2022-11-19 오전 6:00:00  
수정 2022-11-19 오전 6:00:00



김응열 기자

N 기자구독

[이데일리 김응열 기자] 글로벌 산업계의 핵심으로 떠오른 반도체 뉴스가 쏟아지고 있습니다. 우리 결의 가전제품은 낱알이 똑똑해지고 어려운 기술 용어도 뉴스에 속속 등장하고 있습니다. 봐도 봐도 어렵고 알다가도 모르겠는 전자 산업, 그 속 이야기를 알기 쉽게 '톡(Talk)'해드립니다. <편집자주>

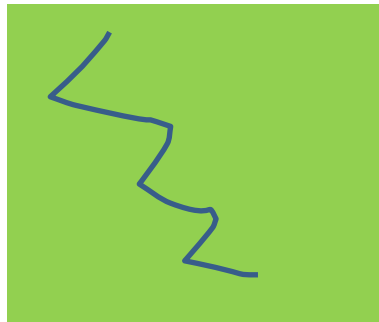
전자·가전기업들이 인공지능(AI) 도입에 속도를 내고 있습니다. 제조 과정에서 인공지능을 적극 활용해 생산성 향상과 품질 개선을 도모하겠다는 계획입니다.

19일 업계에 따르면 삼성전자(005930)는 평택과 기흥, 화성 등 전 반도체 제조공장에서 웨이퍼 이송 로봇(OHT)과 디지털 트윈 시스템 등 AI 시스템을 도입해 활용하고 있습니다.

ta Science Lab & Unique AI

### “Data Annotation”

- 제품 생산의 이미지를 바탕으로 불량 탐지 모델링을 위한 이미지 라벨링

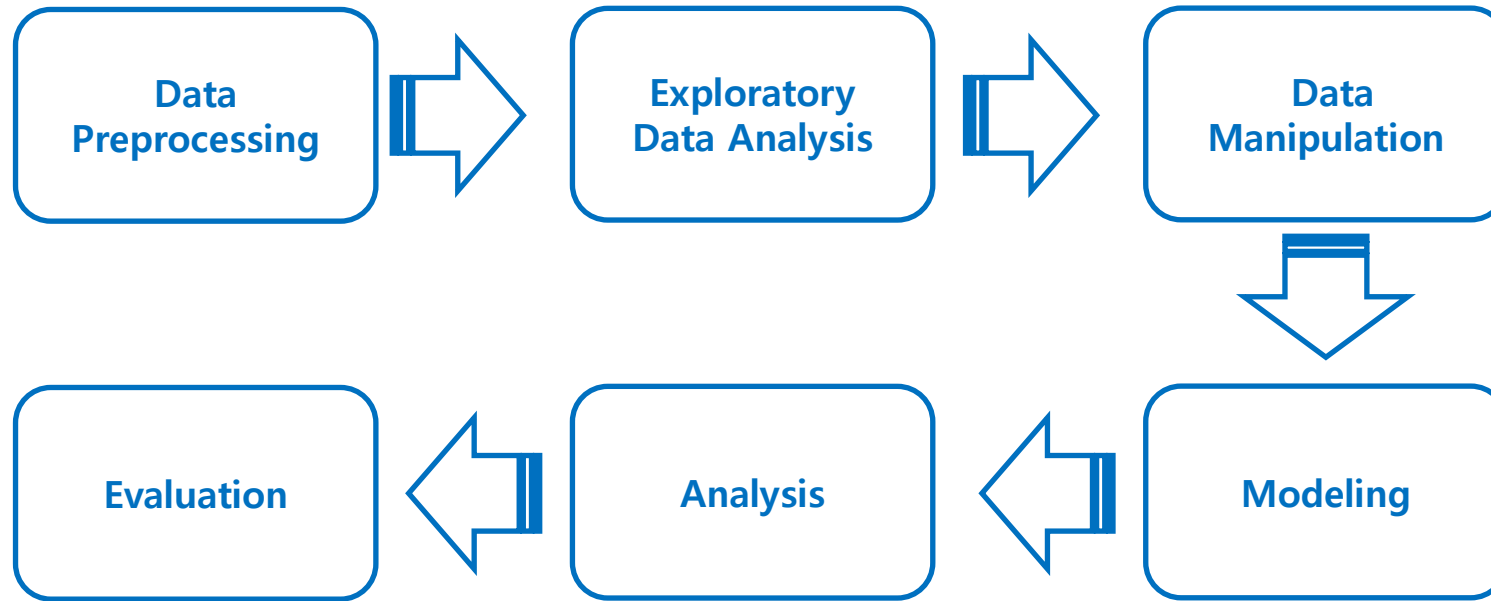


“불량”



“정상”

## II. 품질4.0과 데이터 사이언스



## II. 품질4.0과 데이터 사이언스

### ➤ Data Analytics 모형 구분

#### 지도학습 (Supervised Learning)

종속 및 독립변수를 이용하여 주어진 독립(설명)변수를 바탕으로 종속(반응)변수 예측 모형 제시

예: 회귀/분류 모형

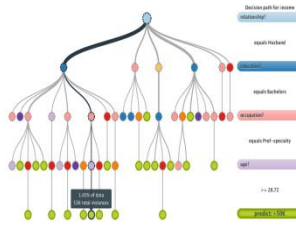


#### 비지도학습 (Unsupervised Learning)

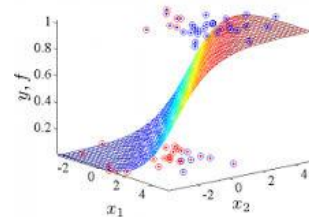
Target(종속변수/반응변수)이 없으며, 독립(설명)변수 간의 관계나 이를 바탕으로 개체들을 구분하여 의미 있는 결과를 제시

예: 군집 분석, 연관성 분석, 주성분 / 요인 분석

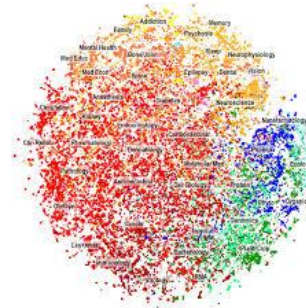
[decision tree]



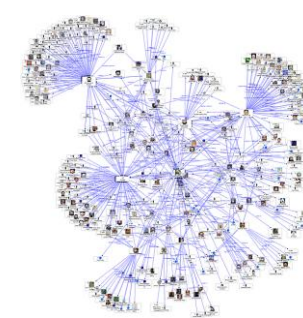
[logistic regression]



[clustering analysis]

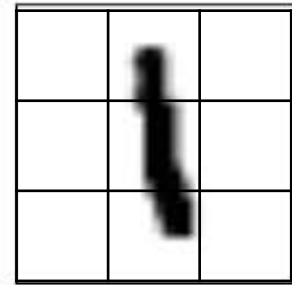


[link analysis]



## II. 품질4.0과 데이터 사이언스

### 데이터



0	254	0
0	254	0
0	254	0

## II. 품질4.0과 데이터 사이언스



<https://www.youtube.com/watch?v=-D2GcuGZReM>



## II. 품질4.0과 데이터 사이언스



<https://www.youtube.com/watch?v=6-DaWgg4zF8>



## II. 품질4.0과 데이터 사이언스

### • 국가별 스마트 팩토리 현황



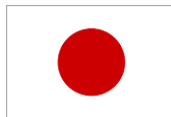
『Industrie 4.0』을 통한 스마트팩토리 전략 추진  
제조설비+IoT+CPS\* ⇒ 맞춤형 유연생산, 제조공정 시뮬레이션 기반 효율 생산  
Industrie 4.0의 표준화 로드맵 v1.0 개발  
참조 구조 모델 (Reference Architecture Model, RAMI4.0) 개발



『첨단제조기술 (Advanced Manufacturing Technology)』 전략 추진  
3D 프린팅 + 디지털 디자인 역량 강화 ⇒ 시제품 제작기간·비용 절감, 스타트업 촉진  
해외 생산기지 구축 (오프쇼어링) ⇒ 자국으로 생산기지 U-턴 (리쇼어링) 전략 전환  
제조혁신을 통한 제조업 효율화 정책, 제조 R&D 강화



『중국제조 2025』 전략 추진  
장비핵심기술 강화, 제품의 품질혁신 ⇒ 독일 Industrie 4.0을 벤치마킹  
소재·부품·공정·산업기술 등 4대 기반 강화를 위한 '공업기반 강화 중장기 계획 발표  
심천을 중심으로 민간 자생적 신기술 제조기반 스타트업의 메카로 부상

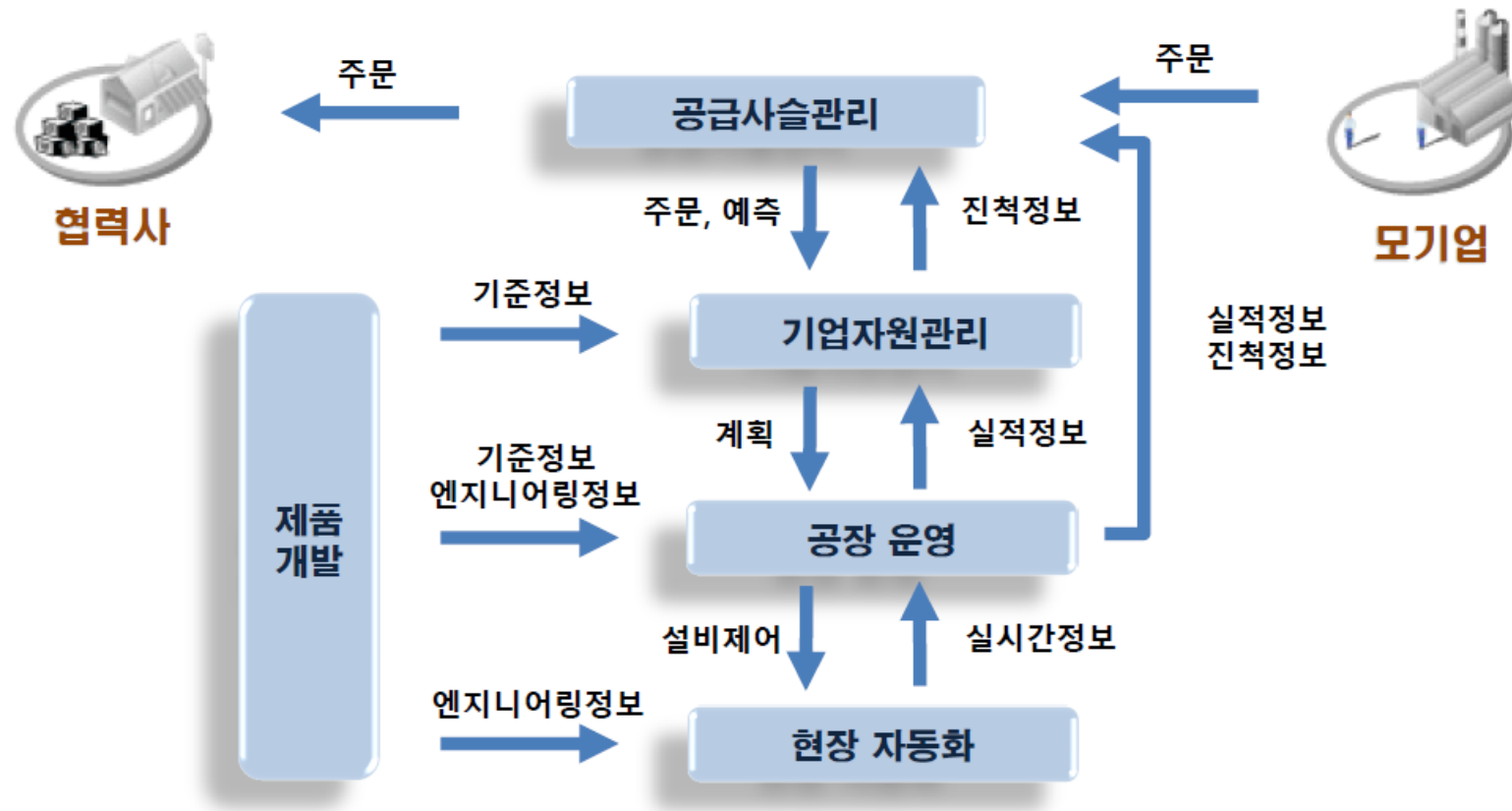


『산업재흥(産業在興)』 플랜을 기반으로 한 산업구조혁신  
범부처 전략적 혁신진흥사업(Strategic Innovation Promotion Program)으로 인프라혁신  
스마트제조의 데이터 모델과 관련된 표준화에 강한 의지

- 4차 산업혁명에 대응하여 제조업의 중요성 기반으로 국가적 노력
- CPS(Cyber Physical System, 사이버물리시스템) : Reactive, Concurrency, Feedback control of physical, Real-time computing이 특징

## II. 품질4.0과 데이터 사이언스

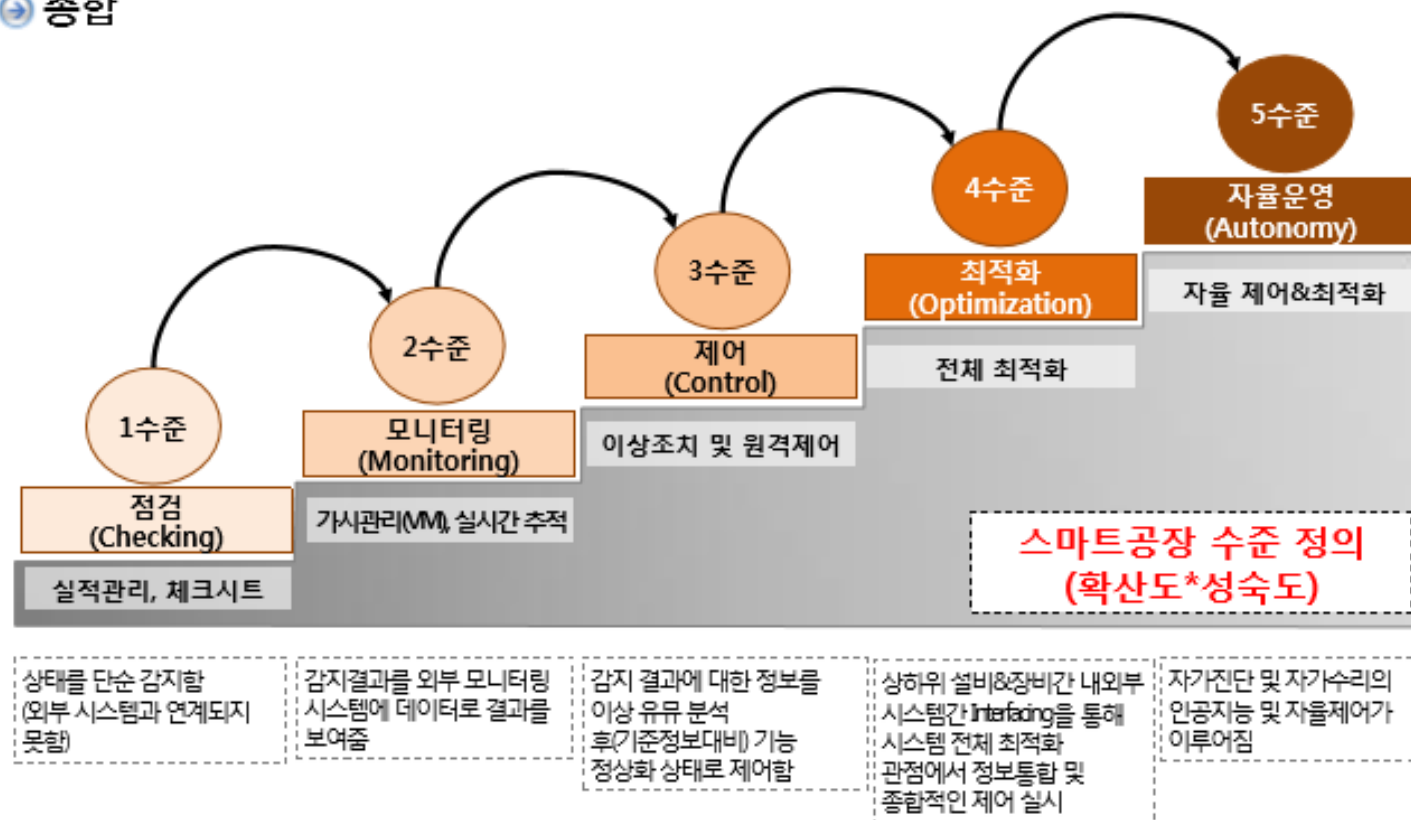
- 산업 내 스마트팩토리



## II. 품질4.0과 데이터 사이언스

- 참고: 스마트 팩토리 성숙도

➡ 종합



\* 모듈별, 세부평가항목별 수준 정의는 유형별로 다르게 정의될 수 있음

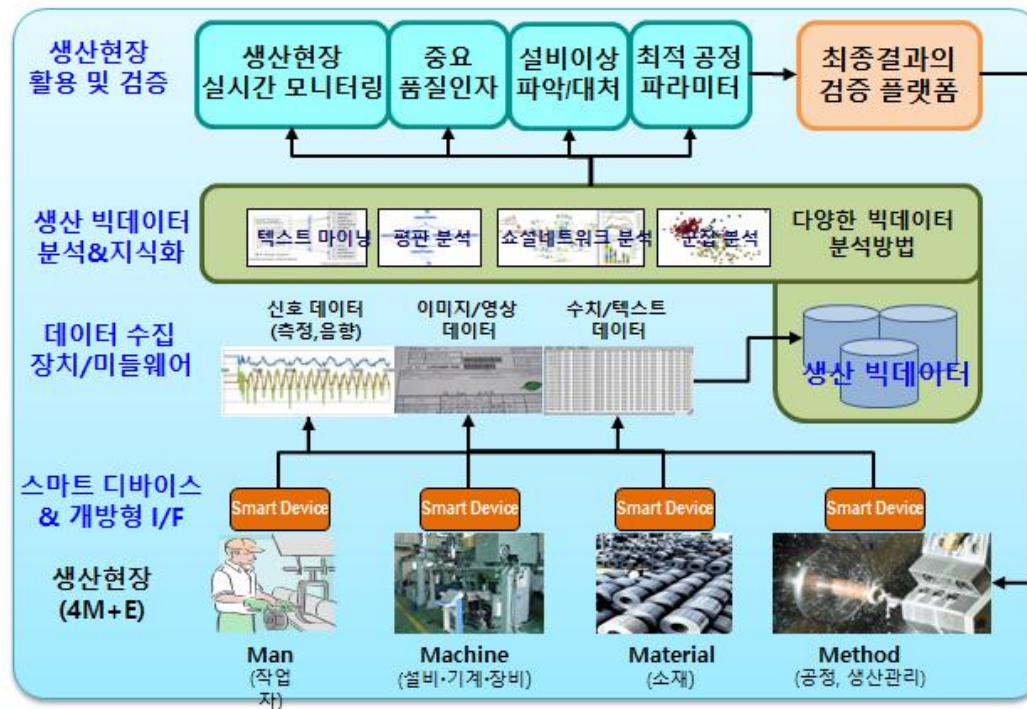
- 스마트팩토리 생산/운영 시스템
  - 생산 · 운영관리의 대상 : 4M
    - Material / Method / Man / Machine
- 생산 및 운영 관리를 통해 달성하고자 하는 기업 목표
  - Product / Quality
  - Cost / Delivery
  - Safety / Morale

- 생산시스템
  - 생산시스템의 설계와 구축
    - 제품생산을 위한 준비과정
    - 불연속적
    - 제품설계, 공정설계, 생산능력 결정, 입지선정, 설비배치, 작업설계, 작업측정 등
  - 생산시스템의 운영
    - 시스템을 운영하여 제품을 생산
    - 연속적이고 계속적

### III. 관련 사례

- 스마트팩토리 사례

- 주조공장의 생산현장에 생산자원(4M1E(Man, Machine, Material, Method, Energy))을 실시간 데이터 취득용 보급형 스마트 디바이스(IoT기반 데이터 수집기와 작업자용 터미널) 개발
- 평균 생산성 10% 이상, 스마트공장 중간 1 수준, 주조기에서 제품 Lot 추적 가능



### III. 관련 사례

- 표준화의 중요성

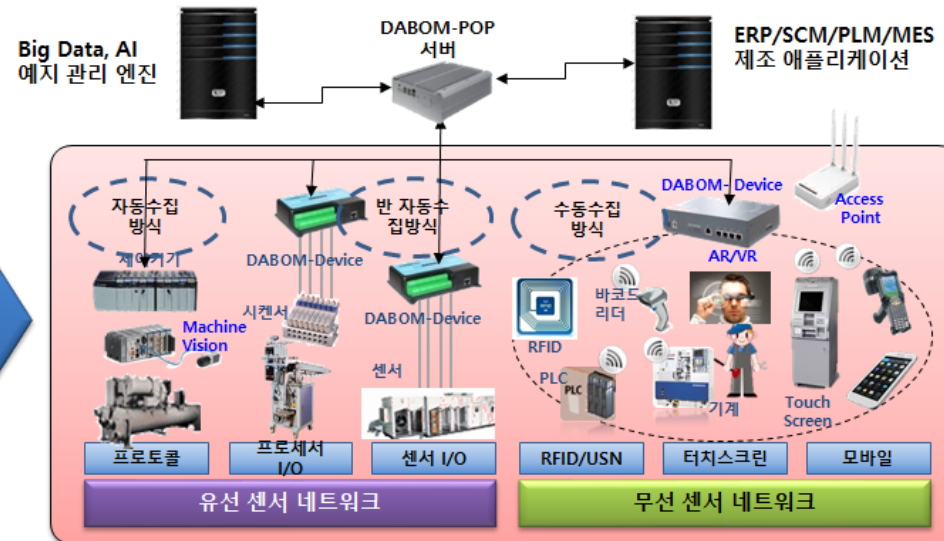
#### 4M1E 생산자원 실시간 통합화

- ISO 22400 KPI for MOM 실시간 생산성과지표 관리 국제표준적용
- 4M1E 생산자원 통합화에 자동, 반자동, 수동 수집 방법과 유선 무선 방식 국제표준 ISO/IEC JTC 1 IoT use case 적용

정보원	내용
 기계	동작횟수 가동시간 고장원인
 설비	전압, 전류 유량, 압력 전력량
 제품	위치, 특성 성능, 치수 제품명
 작업자	Lot번호 시작/종료시간 불량코드 비가동 사유
 에너지	사용량 시간

생산자원 4M1E

ISO 20140  
ISO 22400  
IEC62264  
OPC  
ISO JTC 1 IoT



TTA Standard

정보통신산업표준(국문표준) 제정일: 2017년 6월 28일  
TTAK.KO-11.0227-Part2

생산 자원(4M1E) 기반 스마트팩토리  
정보 관리 - 제2부: 상호 연동지침

Information Management of Smart Factory  
based on Production Resources (4M1E) -  
Part 2 : Interoperability Guideline



TTA 한국정보통신기술협회  
Telecommunications Technology Association



### III. 관련 사례

#### • 발전소 사례

##### – 발전소 운영 효율과 안정성 확보를 위한 활용

#### 발전 산업의 디지털 트랜스포메이션

- 발전소의 최우선 목표는 예방정지와 고장복구이므로 이를 중점적으로 이루어지고 있음
- GE의 Predix는 방대한 양의 발전 운전데이터를 가공, 분석, 예측진단하는 플랫폼 활용
- 국내는 AI기반의 이상탐지, 예방정비, 원격진단, 자산최적화, 디지털 트윈 등 연구개발 단계
- 국내외 공통적으로 비정형데이터의 연구 사례는 부재

<글로벌 기업들의 발전분야 디지털 트랜스포메이션 사례>

기업명	내용
CHINA DATANG	발전소 데이터 분석을 통한 제어, 모니터링, 예지정비, 현장작업자 위치추적, 3D 디지털 트윈을 통한 최적화
TEPCO	IoT 플랫폼을 구축하여 발전소 자산관리를 통해 2017년 10~20% 고장감소성과를 발표
UNIPER	발전설비 오작동 예측으로 운영비 절감 및 오작동 예방
VERMONT	데이터분석을 통해 신재생 에너지 발전량 예측
ENEL	발전설비 청소에 무인항공기와 로봇 활용, 작업자 훈련에 증강현실 활용하여 안전성 향상

- 주로 설비의 온도, 압력, 진동 등 센서 데이터의 분석을 통해 조기경보, 설비의 수명예측 등 활용
- 발전소의 비정형데이터는 수십년간의 지식이 축적된 자산임에도 활용도가 낮음



- 발전소의 디지털트랜스포메이션 전략에 대해 분석
- 비정형 데이터에 대한 데이터 큐레이션이 필요
- 비정형데이터를 활용하여 발전소 운영 효율성의 가치를 높이고자 함

#### 발전소 자산관리 현황

- 발전소 자산관리란 설비의 구매, 운영, 정비, 보수 등 전 주기에 따라 최적의 자산관리를 수행하는 것
- 설비의 주기적인 예방정비를 통해 사전에 고장을 방지하고, 고장발생 시에는 신속한 복구가 필요
- 평균 7,000여개의 설비가 있으며 온도, 압력, 유량 등 초당 약 30만건의 센서데이터 발생
- 국내외에서 빅데이터와 AI를 활용한 설비의 수명예측, 조기경보, 감시 등에 대한 연구가 진행되고 있으며 기술의 성숙도가 실용화 단계

### III. 관련 사례

- DX + 6시그마?

#### 학술세션 IV

#### 인공지능 및 빅데이터 (세미나 7)

#### 좌장 | 이원상 교수 (강릉원주대학교)

- | 예측 인공지능(Predictive AI): 품질경영을 위한 게임체인저  
김계수 (세명대학교)
- | 의약품 공정분석기술 적용을 위한 노코딩 인공지능활용에 관한 연구  
김강희 (LG화학 생명과학사업본부 혁신교육)
- | 설비 고장 진단을 위한 신호 데이터의 대조 학습 기반 과대표집 알고리즘  
신승호, 유재홍 (인천대학교)
- | 제조공정도를 활용한 인공지능 및 빅데이터 활용사례  
최문석 ((주)CE경영컨설팅)

#### 학술세션 IV

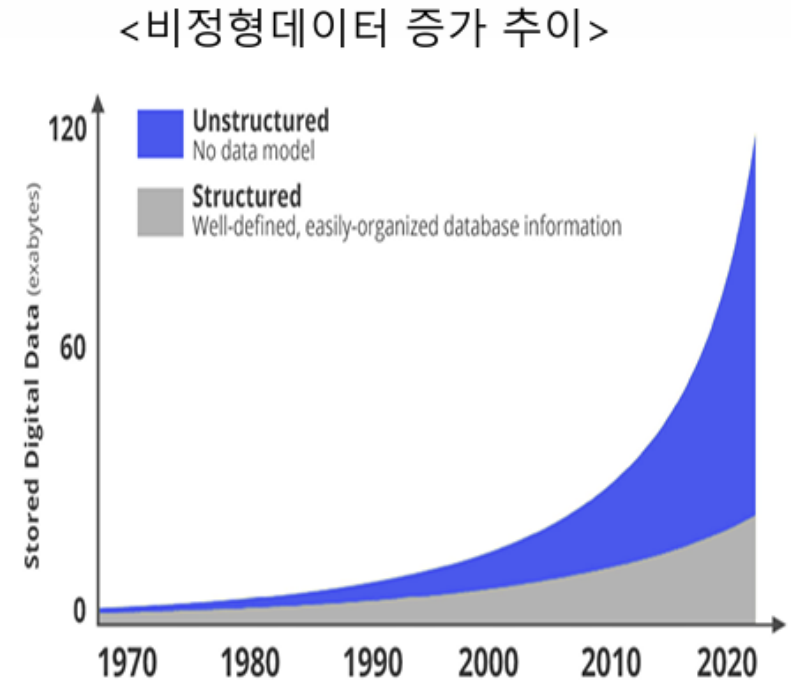
#### LG 화학 Quality Innovation (세미나 5)

#### 좌장 | 유성진 교수 (승실대학교)

- | 석유화학 Digital Twin 구축  
한상훈 (LG화학 여수공장)
- | CNT 중합공정 개선으로 위험지수 감소  
한상훈 (LG화학 여수공장)
- | 5분 멈춤 작업 전 안전점검회의(TBM) 활동 운영  
최영민 (LG화학 여수공장)
- | 밀폐공간 작업 무선통신 가스감지기 모니터링 시스템 구축 및 활용  
최영민 (LG화학 여수공장)

### III. 관련 사례

- 정형+비정형
  - 비정형 데이터의 비중 및 증가 추세
  - 정형데이터와의 상호 보완
  - 정형데이터에서 감지 불가능한 케이스 등의 발견
  - LLM의 활용
    - 예: 작업일지 비정형 데이터 분석/ 사고보고서 분석/특허 분석 등



### III. 관련 사례

모듈	내용	시간	일
제조 및 품질 데이터 사이언스	제조 및 품질과 AX 개요 품질4.0과 데이터 사이언스, 관련 사례	4hr	1일차
제조 데이터 전처리	제조 데이터 전처리: 전처리를 위한 데이터프레임 활용 팁 /결측치 처리/이상값 발견/값 대체/원핫인코딩과 라벨인코딩/ 스케일링/파티셔닝/오버샘플링/이미지읽기/OCR을 통한 이미지2텍스트/IoT센서데이터 처리 등	16hr	2-3일차
제조 데이터 EDA	제조 데이터에 대한 탐색적 데이터 분석 및 해석, 시각화 주요 통계량 학습과 해석/분산분석	8hr	3-4일차
머신러닝 활용 제조 데이터 사이언스 ①	지도학습-분류 기법(Tree 기반 모형과 앙상블 기법 중심)의 이해 및 실습 분류 모형 성능 평가와 파라미터 튜닝을 위한 최적값 탐색 / 예측 수행 지도학습-회귀 (ML 기반 및 선형모형 기반)의 이해 및 실습 회귀 모형 성능 평가와 파라미터 튜닝을 위한 최적값 탐색 / 예측 수행	20hr	4-6일차
통계적 품질 관리(SPC)와 시계열 데이터	통계적 품질관리와 6시그마 개요 시계열 데이터와 SPC, 주요 Control차트 해석 , 시계열 모델링과 시계열 예측, 이상값 탐지	8hr	7일차
머신러닝 활용 제조 데이터 사이언스 ②	업리프트 모델링 이해 및 사례 / 업리프트 트리의 이해 및 실습 / 주요 인과 발견 기법의 이해 및 실습 제조 데이터에 대한 인과 발견 및 그래프 분석과 해석 제품 추천을 위한 기법 : 패턴 발견 및 협업필터링	16hr	8~9일차
자연어 처리와 텍스트 분석	고객 피드백 분석을 위한 자연어 처리 텍스트 분석: 문헌 분류와 감성분석, 토픽모델링 트랜스포머를 이용한 텍스트 분석 및 생성	16hr	10~11 일차
팀 프로젝트	중간 프로젝트를 확장하여 불량 요인 발견 및 모니터링 프로젝트	8hr	12일차



## **Industrial Data Science Lab & Unique AI**

Contact:

won.sang.l@gmail.com

<https://sites.google.com/view/idslab>