

# AI 개발의 혁신견인차, 'MLOps'란?

[애자일] KIAT 산업기술정책센터 정책기획실('21.08.17)

- ◆ 데이터가 범람하는 AI 시대, AI 개발의 효율성·생산성 혁신 동력으로 부상하는 MLOps (Machine Learning Operations)의 개념과 기업 동향을 살펴보고 정리

\* MLOps는 딜로이트가 매년 발표하는 비즈니스 혁신기술 테크트렌드에 2021년 기술로 선정

## < MLOps(ML옵스)란? >

데이터 관리 및 ML 시스템 개발과 서비스 운영(Operations)을 통합해 안정적으로 상품·서비스를 출시할 수 있도록 신속·유연한 개발을 추구하는 협업 방식

- \* MLOps 용어는 구글의 'ML 시스템에 숨겨진 기술 부채(Hidden Technical Debt in Machine Learning Systems)' 보고서('15)에 처음 등장
- \* AIOps는 IT운영 최적화를 위해 사용할 수 있는 AI 기능 스펙트럼을 의미하며 ML 모델 운영에만 초점이 맞춰진 MLOps보다는 상위 개념(가트너, '17)
- 개발, 통합, 테스트, 출시, 배포, 인프라 관리 등 ML 시스템 구성의 모든 단계에서 모델 재학습 자동화 및 모델 모니터링을 지원하며 ML 모델 라이프사이클을 관리

## < 머신러닝 모델 자동화 프로세스 >



출처 : 실리콘밸리의 MLOps, 머신러닝 서비스 구축을 위한 실전 MLOps 가이드(Superb AI, '21)

## 1. MLOps의 부상과 기업 현실

- AI 시장 확대\*로 사진, 동영상, 텍스트 등의 비정형 데이터 양이 증가하며, 데이터 구축을 넘어 데이터 관리 및 활용 차원에서 MLOps의 중요성이 증대\*\*

\* 전 세계 AI 시장규모가 '25년까지 연평균 38.4% 성장해 1,840억 달러(약 203조 6,700억 원) 규모의 시장 형성 전망(BCC Research, '20)

\*\* MLOps의 시장 규모는 '25년 약 40억 달러로 전망(구글 개발자 컨퍼런스 I/O, 가트너, '21)

- 기업들은 최근 비즈니스 문제에 ML을 적용하여 방대한 양의 데이터에서 반복되는 유형을 발견하고 변칙을 파악하여 예측·의사결정과 서비스에 활용하는 경우가 증가

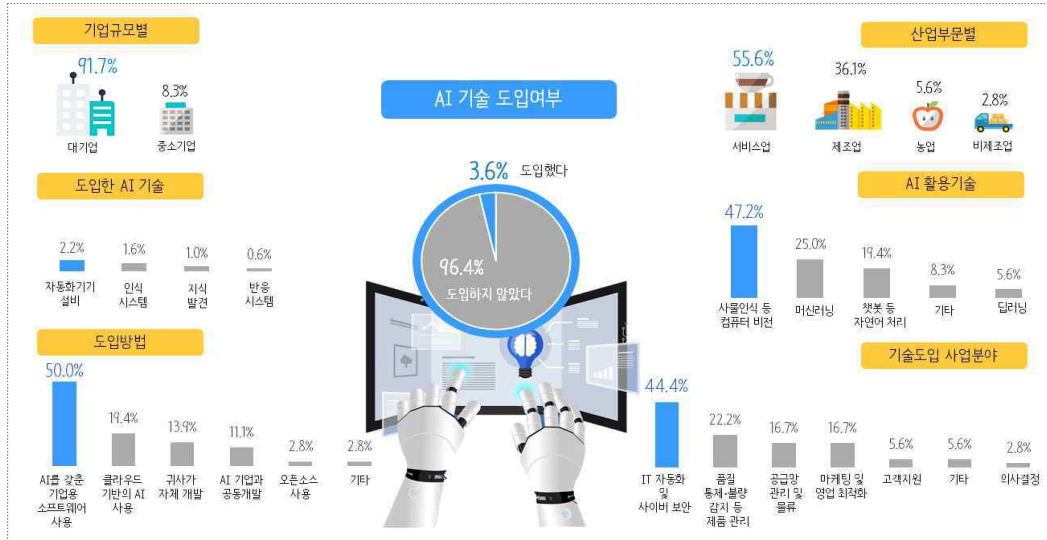
\* 제조시설, 풀필먼트<sup>1)</sup> 센터, 식품가공 공장 등의 산업 환경에 머신러닝 도입하려는 기업 증가

1) Fulfillment, 물류 전문업체가 물건을 판매하려는 업체(셀러)들의 위탁을 받아 보관, 포장, 배송, 재고관리, 교환·환불 서비스 등의 모든 과정을 담당하는 '물류 일괄 대행 서비스'로 국내 기업들은 오픈마켓 경쟁력을 강화하기 위해 서비스 구축

- 이처럼 기업들은 생산성 향상 차원에서 AI, ML 도입의 필요성을 인식하고 있으나 대부분 전문인력 미확보 및 전문인력 확보 후에도 단절적인 조직 문화로 AI 도입이 지체

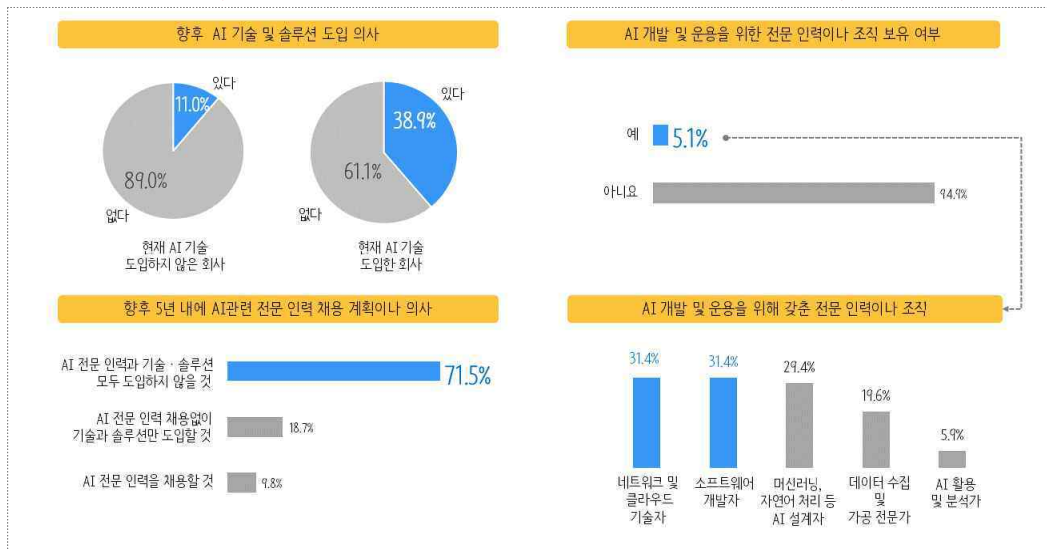
\* 국내 기업 13,225개사 중에서 AI 도입 기업은 409개사에 불과(기업활동조사, '19)

### 〈 국내기업의 AI 도입 현황 〉



출처 : 여론분석-인공지능(AI)(KDI, '20)4

### 〈 국내기업 AI 기술 도입 의사 및 전문 인력 보유 현황 〉



출처 : 여론분석-인공지능(AI)(KDI, '20)

- AI 도입 기업 대부분은 '경영성과 개선'을 경험하였으며 신제품 개발 및 경영관리 측면에서 도움이 되었다고 응답하고 있으나, 향후 AI 기술 도입에는 회의적인 상황

\* AI 기술 도입의 경영 및 성과 평가(% , '20) : 도움(77.8), 보통(19.4) 도움이 되지 않았다(2.8)

\* AI 기술 도입을 통한 성과 분야(% , '20) : 신제품 개발(32.1), 경영관리(25.0), 비용·생산관리(25.0), 인력·노무 관리(17.9)

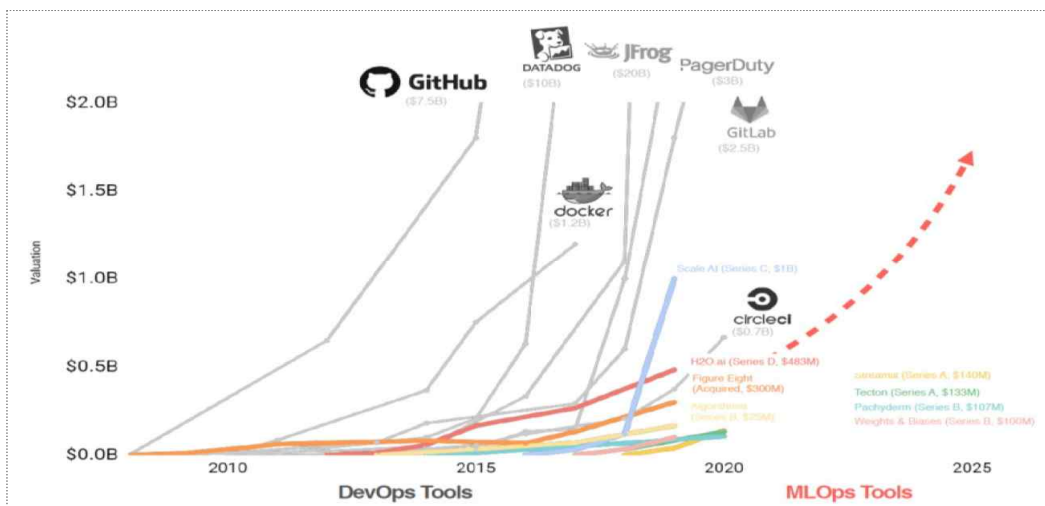
- 전문인력 부족 및 전문조직(부서)의 부재, 기업 수요에 적합한 솔루션 부족, 조직 내부의 의사결정 시스템 등의 이유로 기업들은 AI (추가)도입을 신중하게 검토
  - \* AI 도입 및 활용의 가장 큰 애로 요인으로 '적합한 기술을 보유한 인력 고용의 어려움'으로 나타남 ('기업의 AI활용 실태 조사', 산업연구원, '21)
- 결국, AI 기술 도입 기업들은 AI 기술 개발보다는 'AI를 갖춘 기업용 SW(50.0%)를 주로 사용 중이며, 기업 수요에 적합한 솔루션 개발과 기업 내부에서의 안정적인 AI 운용 시스템을 희망
  - \* '머신러닝'(25.0%), '딥러닝'(5.6%) 등 원천 기술보다 '사물인식 등 컴퓨터 비전'(47.2%)과 같은 완성형 기술을 가장 많이 활용
- 글로벌 대기업들은 AI 운용 시스템으로써 DevOps(데브옵스), MLOps 등을 자체 개발 하고 있으며, 기업용 AI 솔루션 개발 스타트업도 성장하고 있는 상황
  - \* MLOps는 ML의 산업화와 확장을 위한 개발 및 전달 과정에 DevOps(데브옵스)의 도구 및 접근 방식을 적용하는 것으로 데브옵스와 유사한 형태로 발전 중

### 〈 DevOps란? 〉

- (정의) 개발(Development)과 서비스 운영(Operation)을 통합하는 협업 방안
- (배경) 소프트웨어 개발팀과 서비스 운영팀 간 정보가 유연하게 교류되지 못하면 시장의 변화나 고객의 니즈를 빠르게 제품에 반영하지 못하는 비효율 존재
- (대표기업) 깃허브(GitHub, 코드 관리 플랫폼), 데이터 독스(Datadog, 서비스 운영 인프라 관리)

구분	DevOps		MLOps
개념	Development + Operation	↔	Machine Learning + Operation
주요 시기	2000년대, 소프트웨어 시대		2010년 이후, 인공지능 시대
개발 원자재	코드		코드, 데이터 *데이터 관리와 지속적인 학습이 필수적

### 〈 DevOps와 MLOps 기업 성장 추세 〉

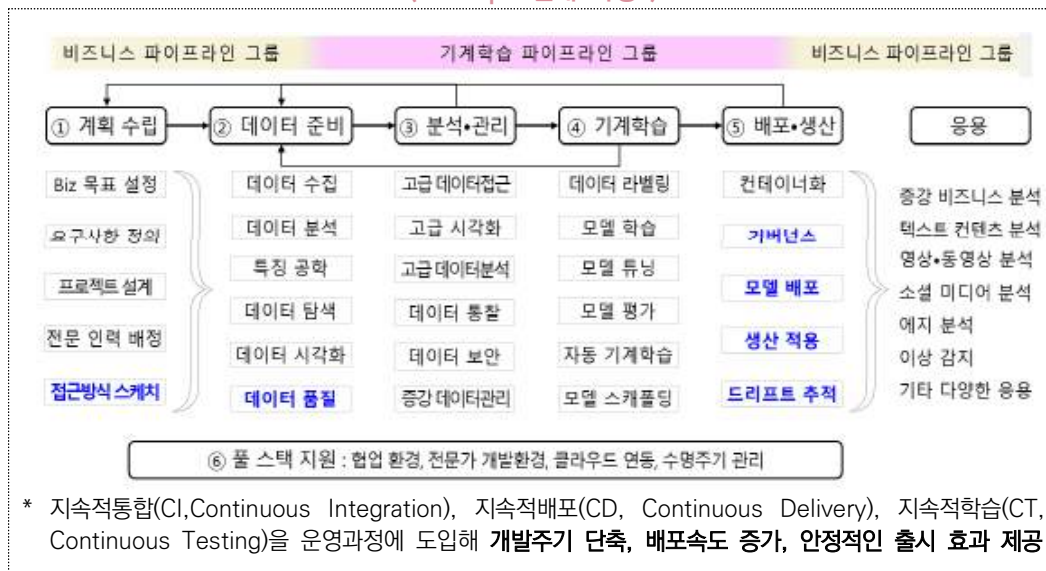


출처 : 기사“MLOps 관점서 데이터 바라봐야”, DATANET('20)

## 2. MLOps의 중요성

- ML 프로젝트는 생산 단계에서 발견된 심각한 문제로 전체 프로젝트를 재설계하거나 학습 모델을 폐기하는 경우가 다수이고 이외에 부서 간 이해 갈등 등 **다양한 장애 요인이 존재**
  - 고정된 데이터를 통해 최고의 성능을 가진 모델을 도출하고자 하는 **ML 모델 개발자**와 계속 해서 변하는 데이터를 토대로 안정적인 운영이 필요한 **서비스 관리자** 사이에 간극 발생
  - 데이터, 개발, 운영이 분리될 경우 데이터 사일로가 발생하면서 급변하는 시장 상황에 대응이 느려지고 AI 성능 저하문제가 발생하며, 나아가 서비스 운영에도 비효율 초래
- ML 프로젝트의 숨겨진 기술 부채(Hidden technical debt)<sup>2)</sup>를 통제할 수 있는 기존과 다른 형태의 운영 시스템이 요구되며 MLOps 필요성 증대
  - ML 프로젝트를 단순 모델설계로 간주할 경우, 모델의 비즈니스로 활용 과정 등 이를 제외한 나머지가 숨겨진 기술 부채\*로 작용하여 프로젝트 실패의 가능성이 증가
    - \* 기계학습 모델의 일부 수정 불가능, 학습과 생산 데이터가 일치하지 않을 때의 예측 불가능성, 입력과 출력 사이의 블랙박스인 해설 불가능성 등
  - ML 프로젝트의 작업 흐름을 매끄럽게 도와줄 수 있는 도구와 소통의 협업 방식으로 생산성을 개선시킬 수 있는 MLOps의 중요성이 증가

### 〈 MLOps 전개 과정 〉






출처 : 기계학습 운영화(MLOps) 실행 전략(한국전자통신연구원, 조성익·유용식·표철식, '21.2.24)

- 2) 기술 부채는 현 시점에서 더 오래 소요될 수 있는 더 나은 접근방식을 사용하는 대신 쉬운(제한된) 솔루션을 채택함으로써 발생하는 추가적인 재작업의 비용을 반영하는 소프트웨어 개발의 한 관점 ex) 개발 일정을 맞추기 위해 코드 검토를 소홀히 하는 것, 반복되는 일(배포 등)을 자동화 하지 않는 것 등

### 3. MLOps 플랫폼 기업동향

- MLOps 생태계는 데이터 플랫폼을 중심으로 모델 배포/관리, 분석, 학습, 라벨링 등의 영역에서 특화된 서비스들이 생겨나며 성장하고 서로 통합되며 발전을 거듭할 것으로 예상
- (대표 기업) MS, AWS, 구글 등 글로벌 주요 IT기업들은 복잡한 절차가 필요한 AI·ML 기술개발의 효율성과 생산성을 향상하고자 MLOps 서비스를 개발하여 지원

#### 〈 MLOps 플랫폼 대표 기업 〉

기업	MLOps 서비스 내용
 Microsoft	<ul style="list-style-type: none"> <li>• 마이크로소프트는 클라우드 기반의 '애저 머신러닝 서비스(Azure ML Service)'를 통해 개발자가 맞춤형 ML 모델을 손쉽게 개발·운영 할 수 있도록 지원</li> </ul>
	<ul style="list-style-type: none"> <li>• 아마존웹서비스(AWS)는 '세이지메이커(SageMaker)'를 통해 전체 ML 워크플로우를 포괄해 작업을 수행하는 완전 관리형 서비스 제공</li> <li>• AWS는 ML서비스를 통해 제조업의 생산성과 품질 제고 실현을 목표로 추진</li> </ul>
 Google Cloud	<ul style="list-style-type: none"> <li>• 구글 클라우드는 'AI 플랫폼(AI Platform)'에서 데이터 라벨링, 노트북, 대시보드, AI 허브 등 머신러닝 라이프 사이클 관리를 지원하는 다양한 기능을 제공</li> </ul>

출처 : 보도자료 및 보고서 자료 정리

- (스타트업) 셀던 등 스타트업에서도 ML 라이프사이클 관리를 돕는 MLOps 지원 서비스 개발

#### 〈 MLOps 플랫폼 스타트업 〉

기업	MLOps 서비스 내용
알고리즘미아 (Algorithmia)	<ul style="list-style-type: none"> <li>• ML 포트폴리오에 연결되어 배포, 관리, 확장 기능을 제공</li> </ul>
셀던 (Seldon)	<ul style="list-style-type: none"> <li>• 엔터프라이즈 기능을 추가로 제공하는 오픈 코어 서비스 기업으로 ML 모델 점검 및 설명을 위한 오픈소스 알리바이(Alibi) 라이브러리 제공</li> </ul>
W&B (Weights & Biases)	<ul style="list-style-type: none"> <li>• 종합 실험 추적 서비스 제공, ML 라이브러리 즉시 통합, 모델이 개발에서 생산 과정으로 가는 과정에서 어떤 입력값과 변화에 대한 감사 흔적 제공</li> </ul>
메타플로우 (Metaflow)	<ul style="list-style-type: none"> <li>• ML 라이프사이클 관리에 특화된 파이썬 친화적인 코드 기반 워크플로우 시스템으로 워크프로우 설계, 프로덕션 배포에 유용</li> </ul>
데이터로봇 (Datarobot)	<ul style="list-style-type: none"> <li>• 데이터 준비에서 생산 배치에 이르는 사용자의 엔터프라이즈 AI 생애 주기를 모두 관할하려는 기업</li> </ul>
페이퍼스페이스 (Paperspace)	<ul style="list-style-type: none"> <li>• 데이터를 탐색하고 신경망을 학습시키고 프로덕션급 ML 파이프라인을 구축하기 위한 툴모음</li> </ul>
파키덤 (Pachyderm)	<ul style="list-style-type: none"> <li>• AI/ML용 강력한 데이터 버전 관리 및 데이터 플랫폼을 제공하여 데이터 변환, 훈련 및 모델 개발을 빠르게 확장</li> </ul>
ML플로우 (MLflow)	<ul style="list-style-type: none"> <li>• 오픈소스 ML 라이프 사이클 관리 플랫폼으로 ML플로우 추적을 사용하면 AP 호출을 사용해 실험 기록 및 쿼리 가능</li> </ul>

출처 : 보도자료 및 보고서 자료 정리



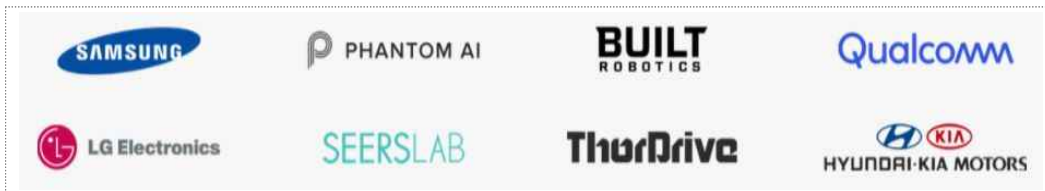
- (기업연합) AI 개발의 효율성 극대화 및 MLOps 저변 확대를 위한 글로벌 AI 기업 연합 'AI 인프라스트럭처 얼라이언스(AI Infrastructure Alliance)' 출범('19)
  - 미국과 영국, 독일 등 세계 각국에서 뛰어난 기술력을 높이 평가 받고 있는 스타트업\*을 시작으로 국내 기업 중에서는 유일하게 슈퍼브에이아이(Superb AI)\*가 합류('20.10)
  - \* (한국) 인공지능 데이터 플랫폼 기업 Superb AI는 데이터라벨링 등의 AI 데이터 전처리 작업을 포함해 데이터 구축, 관리, 분석 등의 데이터 처리 작업을 통합적으로 지원하는 올인원 플랫폼 '스위트3' 운영

#### 〈 AI Infrastructure Alliance 기업 명단 〉



출처 : Superb AI

#### 〈 Superb AI 파트너사 〉



출처 : Superb AI

## 4. 시사점

- MLOps는 안정된 인프라와 자동 운영 시스템을 제공하여 ML 모델 개발과 운영에서 반복 발생하는 문제를 최소화함으로써 **기업의 비즈니스 가치 창출**에 기여
  - 유통, 제조, 금융, 통신, 의료 등 다양한 산업에서 ML 모델 운영·서비스화를 지원
  - (Uber Eats) MLOps를 활용해 배달 소요 시간 예측 모델을 효율적으로 운영
- 풍부한 데이터를 기반으로 다양한 기술이 적용되어야 하는 AI 기술의 특성을 고려할 때, 방대한 양의 데이터를 효과적으로 처리가 가능한 **ML의 중요성이 증가할 것으로 기대**
  - \* AI 개발 전과정에서 데이터를 가공하고 관리하는 작업이 차지하는 비중은 80% 이상

3) 스위트(Suite)의 대표적인 자동화 기능 중 하나인 '오토라벨링'은 AI가 1차로 데이터 라벨링(원천 데이터에 값을 붙이는 작업)을 진행하고 2차로 AI가 검수를 요청한 부분만 수정하도록 해 수동작업 대비 약 10배 빠른 속도로 데이터 처리

- 데이터 구축이 끝이 아니라 데이터를 잘 관리하고 활용하는 것이 중요한 만큼 신속하고 유연한 AI 개발을 뒷받침하는 **MLOps 관점**에서의 대응이 보다 중요
- 글로벌 IT기업 뿐 아니라 미국, 유럽 등의 주요 국가에서 MLOps 서비스를 제공하는 스타트업이 등장하므로 우리나라도 **MLOps 관련 스타트업 육성 및 창업 지원 활성화 필요**
  - 우리나라는 유통 대기업 중심으로 ML 사용 사례가 증가하는 중이나 MLOps 서비스를 제공하는 기업은 소수에 불과하므로 **기업 수요에 적합한 솔루션 개발** 지원 필요
  - 국내 스타트업이 세계적 흐름에 맞춰 성장할 수 있도록 AI기술·개발현황·운영 노하우 등을 공유 할 수 있는 **기업 간 네트워크 및 연합 체제 구축**을 통해 생태계 형성 지원
- ML 프로젝트 진행에 관여하는 여러 이해관계자들\*에 대한 **실무교육 및 전문인력 양성 필요**
  - \* 데이터 사이언티스트, 데이터 엔지니어, 머신러닝 엔지니어, 데브옵스 엔지니어, IT, 비즈니스 책임자, 관리자 등
  - 우리나라 기업들은 AI 관련 전문인력 확보에 큰 어려움을 겪고 있는 상황이며, 이에 따라 디지털 전환, AI·ML 기술 도입 등이 늦춰지고 있으므로 **전문인력 양성이 절실**
  - AI 전문인력이 많더라도 기업 전체가 디지털 전환이 되지 않거나 경영진과 엔지니어의 인식 문제, 조직의 경직된 문화로 인하여 ML 프로젝트가 실패하므로 **기업 내부 환경 개선도 필요**

#### [ 참고자료 ]

- TechDaily(김태훈), AI 시장 패러다임 바꾸는 'MLOps'.....AI 기술 개발 효율성 및생산성 향상('20.10.15)
- InfoWorld, 머신러닝 운영과 관리, 'MLOps'가 뜬다
- DATANET(강석오), [주목2021]MLOps, AI개발 효율성·생산성 혁신 동력으로 부상('21.01.20)
- InfoWorld, 'MLOps 분야 이끈다' 알아둘 만한 AI 스타트업 5곳
- 이뉴스투데이(고선호), 스타트업계, 인공지능 '동맹' 확대...“B2B 시장 이끈다”(‘21.04.11)
- DATANET(강석오), “MLOps 관점서 데이터 바라봐야”(‘20.12.21)
- ITWorld Korea, 머신러닝 라이프사이클 관리를 돕는 10가지 MLOps 플랫폼
- 슈퍼브에이아이, 실리콘밸리의 MLOps, 머신러닝 서비스 구축을 위한 실전 MLOps 가이드('21)
- IT & Future Strategy 제 5호, 인공지능 플랫폼 글로벌 선도기업 분석('21. 5. 26)
- MLOps: 머신러닝의 지속적 배포 및 자동화 파이프라인, <https://smilegate.ai/2021/06/10/mlops/> ('21.6.10)
- 마소캠퍼스(김진·최정아), 데이터 사이언스 입문
- XBrain(Jongmin), “MLOps란 무엇일까?”, “기계학습을 위한 데이터 전략”
- 인더스트리뉴스(최정훈), AWS 윤석찬 수석테크에반젤리스트, “제조업 숙원 ‘예지정비’ 아마존 머신러닝이 궤도에 올린다”(‘21.2.13)
- AI타임스(박성은), “개발자들 주목하는 MLOps, 지속가능한 AI 서비스 만든다”(‘21.7.7)
- KDI 경제정보센터 여론분석팀, 2020-03 인공지능편 여론분석('20)
- 한국전자통신연구원(조성익·유용식·표철식), 기계학습 운영화(MLOps) 실행 전략('21.2.24)

※ 한국산업기술진흥원 산업기술정책센터 정책기획실 손다슬 연구원 / three@kiat.or.kr

※ 본 자료에 수록된 내용은 작성자의 개인의견으로 기관의 공식 견해가 아님을 밝힙니다.