

# 프로덕션 레디 AI/ML 환경을 구축하기 위해 고려해야 할 중요 사항



# 데이터의 비즈니스 가치 확대

## 중요 비즈니스 자산으로서의 데이터

전 세계적으로 저장된 데이터의 양은 2024년까지 8.9제타바이트에 육박할 것으로 예측되고 있습니다.<sup>1</sup> 디지털 세계에서는 데이터를 사용해 경쟁 우위를 확보할 수 있습니다. 하지만 데이터 수집은 그저 시작일 뿐이며, 진정한 차별화 요소는 바로 데이터를 사용하는 방식입니다.

인공지능(AI), 머신 러닝(ML), 딥 러닝(DL)은 데이터를 사용해 비즈니스 인사이트를 제공하고, 태스크를 자동화하며, 시스템 기능을 발전시킵니다. 이러한 기술은 고객과 직원에서부터 개발과 운영에 이르는 비즈니스의 모든 측면을 혁신할 잠재력을 가지고 있습니다. 소프트웨어 애플리케이션에 AI/ML을 구축하면 측정 가능한 비즈니스 성과를 달성하는 데 도움이 됩니다.

- 고객 만족도 증가
- 차별화된 디지털 서비스 제공
- 기존 비즈니스 서비스 최적화
- 비즈니스 운영 자동화
- 수익 증가
- 비용 절감

## 데이터를 비즈니스 자산으로 전환

본 e-book은 실행 가능한 데이터 분석을 위한 다양한 기술을 설명합니다.

- **인공지능**: 인간의 행동을 모방해 일반적으로 인간의 개입이 필요한 태스크를 수행합니다.
- **머신 러닝**: AI에 속하며 알고리즘과 통계 모델을 사용해 명시적 지침 없이 태스크를 수행합니다.
- **딥 러닝**: ML에 속하며, 인간의 뇌와 비슷한 방식으로 레이어를 사용해 원시 입력으로부터 고수준의 특징을 계속해서 추출해냅니다.

실생활 AI에 대한 핵심 가이드(An executive's guide to real-world AI)를 읽고 AI와 ML의 비즈니스 측면에 대해 자세히 알아보세요.

## 산업 전반의 AI/ML 활용 사례



### 의료

- 임상 효율성 증대
- 진단 속도 및 정확성 향상
- 환자 결과 개선



### 금융 서비스

- 맞춤형 고객 서비스 제공
- 리스크 분석 개선
- 사기 및 자금 세탁 감지



### 통신

- 고객 행동에 대한 인사이트 확보
- 고객 경험 향상
- 5G 네트워크 성능 최적화



### 자동차

- 자율 주행 지원
- 유지 관리 요구 사항 예측
- 공급망 개선



### 보험

- 보험금 청구 처리 자동화
- 사용자 기반 보험 서비스 제공

<sup>1</sup> IDC, “IDC의 Global StorageSphere, 세계의 스토리지 용량 설치 기반이 지속적으로 견조한 성장세를 보일 것으로 예측(IDC's Global StorageSphere Forecast Shows Continued Strong Growth in the World's Installed Base of Storage Capacity)”, 2020년 5월 13일. [idc.com/getdoc.jsp?containerId=prUS46303920](http://idc.com/getdoc.jsp?containerId=prUS46303920)



# 프로덕션 레디 AI/ML 환경 구축

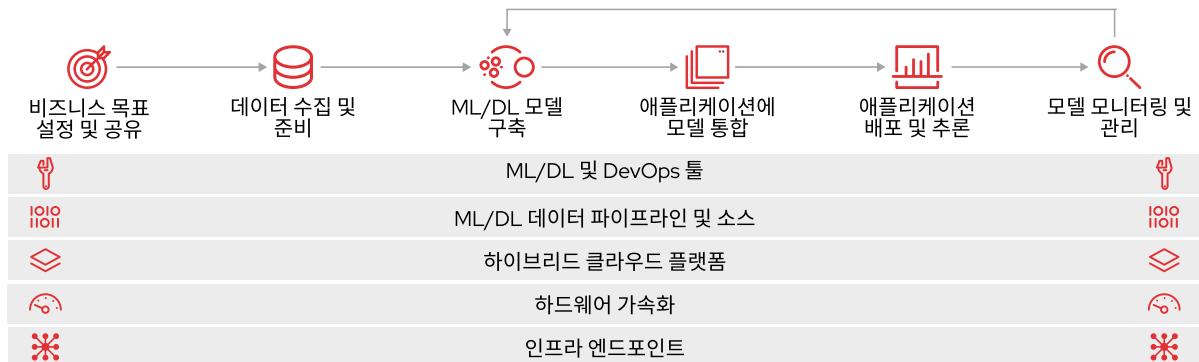
프로덕션에 AI/ML을 배포하는 것은 단순한 AI/ML 모델 생성에서 확장된 반복적 프로세스입니다. AI/ML 라이프사이클의 주요 단계는 다음과 같습니다.

- AI/ML 이니셔티브를 위한 비즈니스 목표를 설정하고, 모든 이해관계자와 이를 공유
- AI/ML 이니셔티브에 필요한 데이터 수집 및 준비
- 목표에 따라 ML/DL 모델 개발
- 애플리케이션 개발 프로세스에 ML/DL 모델 배포
- ML/DL로 구동되는 지능형 애플리케이션을 구현하고 추론 시작
- 시간 경과에 따른 정확성 확보를 위해 모델 모니터링 및 관리

적응형 오픈 AI/ML 아키텍처를 사용하면 이러한 프로세스를 보다 효과적으로 실행할 수 있습니다. 이러한 아키텍처에는 여러 가지 핵심 기술과 기능이 필요합니다.

- AI/ML 및 DevOps 툴**은 데이터 과학자, ML 엔지니어, 애플리케이션 개발자가 ML/DL 모델과 애플리케이션을 구축, 배포 및 관리할 수 있도록 해줍니다.
- 데이터 파이프라인**은 정리된 데이터를 제공해 데이터 과학자는 ML/DL 모델을, 애플리케이션 개발자는 데이터 관리 요건을 생성, 학습 및 검증하도록 합니다.
- 클라우드 플랫폼**은 데이터 엔지니어, 데이터 과학자, ML 엔지니어, 애플리케이션 개발자가 작업에 필요한 리소스에 신속하게 액세스할 수 있도록 지원합니다.
- 컴퓨팅, 스토리지 및 네트워크 액셀러레이터**는 데이터 준비, 모델 개발, 추론 태스크를 가속화합니다.
- 인프라 엔드포인트**는 AI/ML 운영의 모든 단계에서 온사이트, 가상, 엣지, 프라이빗, 퍼블릭, 하이브리드 클라우드 환경 전반에 리소스를 제공합니다.

본 e-book은 효과적인 AI/ML 아키텍처를 구축하기 위한 주요 고려 사항을 검토합니다.



## AI/ML 배포 과제

조직에서는 AI/ML 환경을 구축할 때 여러 가지 어려움을 겪습니다.

- 인재 부족:** AI/ML 전문 인력이 한정되어 있기 때문에, AI/ML 담당자를 찾아 유지하기가 더욱 어렵습니다.
- 손쉽게 사용할 수 있는 데이터 부족:** 기업은 대규모로 데이터를 수집하지만, 각 AI/ML 이니셔티브에 적합한 데이터를 찾고 준비해야 합니다.
- 분산된 팀:** 느리고, 수동적이며, 분산된 운영으로 인해 AI/ML 배포가 지연될 수 있습니다.
- 리소스 제공 지연:** 인프라 및 툴 제공이 지연되면 모델 개발, 통합, 애플리케이션 내 배포에 문제가 생깁니다.

하지만 AI/ML 라이프사이클에 클라우드 네이티브 애플리케이션 개발 접근 방식을 적용하면 이러한 문제들을 극복할 수 있습니다.

# 컨테이너 및 컨테이너 오케스트레이션

## 컨테이너

컨테이너란 종속성을 모두 포함하여 애플리케이션을 패키지화한 소프트웨어의 기본 단위입니다. 컨테이너는 애플리케이션 구축 프로세스를 간소화하고, 애플리케이션을 변경하지 않고도 서로 다른 환경에 배포할 수 있게 해줍니다.

## AI/ML에 중요한 이유

데이터 과학자, ML 엔지니어, 애플리케이션 개발자가 높은 생산성을 실현하려면 본인이 선호하는 툴과 리소스에 액세스할 수 있어야 합니다. 이와 동시에 IT 운영 팀은 이러한 리소스가 최신이며, 컴플라이언스를 준수하고, 안전하게 사용할 수 있는지 확인해야 합니다. 컨테이너는 하이브리드 환경 전반에서 일관된 방식으로 다양한 AI/ML 툴을 빠르고 손쉽게 배포할 수 있게 해줍니다. 팀은 투명성을 위해 변경 사항을 추적하는 버전 관리 기능과 함께 반복적으로 컨테이너 이미지를 수정하고 공유할 수 있습니다. 동시에 프로세스 격리 및 리소스 제어를 통해 위협에 대한 보호 수준을 높입니다.

## 모범 사례 및 권장 사항

통합 보안 기능을 포함하고 있으며 환경 전반에서 컨테이너를 배포, 관리 및 이동할 수 있게 해주는 강력한 고가용성 컨테이너 플랫폼을 찾으세요. 더 유연하고 다양하게 선택할 수 있도록 광범위한 기술을 통합하는 오픈소스 플랫폼을 선택합니다.

## ExxonMobil

ExxonMobil은 오픈소스 컨테이너 플랫폼을 사용하여

## 개선된 협업 환경을

데이터 과학 팀에 제공합니다.<sup>2</sup>

## 컨테이너 오케스트레이션

컨테이너 오케스트레이션은 환경 전반에서 컨테이너의 생성, 배포, 라이프사이클을 관리하는 작업을 포함합니다.

## AI/ML에 중요한 이유

컨테이너를 도입한 후에는 효율적으로 배포하고 관리하고 확장할 방법이 필요합니다. 컨테이너 오케스트레이션 툴은 컨테이너의 라이프사이클을 일관된 방식으로 관리할 수 있게 해줍니다. 이러한 툴은 일반적으로 온사이트, 엣지 및 클라우드 환경 전반에서 컴퓨팅, 스토리지 및 네트워킹 리소스에 대한 액세스를 중앙화합니다. 또한, 통합된 워크로드 스케줄링, 멀티테넌시 제어, 할당량 적용도 제공합니다.

## 모범 사례 및 권장 사항

**쿠버네티스** 기반 컨테이너 오케스트레이션 툴을 선택하여 선도적인 오픈소스 기술을 활용하세요.



## 쿠버네티스

선도적인 컨테이너 오케스트레이션 프레임워크입니다.<sup>3</sup>

<sup>2</sup> Red Hat 고객 사례 “[신속한 정보 공유로 민첩성과 생산성을 확보한 ExxonMobil](#),” 2020년 6월 12일 액세스

<sup>3</sup> Red Hat, “[하이브리드 클라우드, 엔터프라이즈급 쿠버네티스](#),” 2020년 6월 12일 액세스



# 애플리케이션 관리 및 DevOps

## 애플리케이션 라이프사이클 관리

애플리케이션 라이프사이클 관리에는 컨테이너에서 실행되는 애플리케이션의 배포, 확장, 관리가 포함됩니다.

## AI/ML에 중요한 이유

Red Hat 고객의 29%는 진화하는 소프트웨어 스택의 호환성과 복잡성 관리가 AI/ML 이니셔티브에서 가장 큰 과제라고 밝혔습니다.<sup>4</sup> 컨테이너 애플리케이션 라이프사이클 관리 구성 요소는 컨테이너 오케스트레이션 툴에 통합되어 AI/ML 개발 툴을 비롯해 컨테이너화된 애플리케이션을 직접적으로 관리할 수 있게 해줍니다. IT 운영 팀은 설정, 프로비저닝, 업데이트와 같은 일반적인 라이프사이클 관리 태스크를 자동화해 효율성, 속도, 정확성을 확보할 수 있습니다. 데이터 과학자, ML 엔지니어, 애플리케이션 개발자는 IT 운영 팀의 개입 없이 사전 승인된 서비스 카탈로그의 툴과 애플리케이션을 배포할 수 있습니다. 또한, 자동화를 통해 반복적인 태스크를 수행할 필요가 없어진 직원들은 보다 혁신 전략 활동에 더 많은 시간을 투자할 수 있습니다.

## 모범 사례 및 권장 사항

사용이 간편한 자동화 기능과 선호하는 AI/ML 툴과의 통합 기능을 갖춘 컨테이너 애플리케이션 라이프사이클 관리 툴을 선택하세요. 널리 사용되는 툴에는 **쿠버네티스 오퍼레이터**와 **Helm 차트**가 있습니다.



“쿠버네티스 오퍼레이터를 통한 Seldon의 통합은 [...] 조직에서 하이브리드 클라우드 전반에 머신 러닝 모델을 신속하게 배포할 수 있도록 지원하므로 AI 기반 디지털 서비스를 빠르게 롤아웃할 수 있습니다.”<sup>5</sup>

Alex Housley  
Seldon, 창립자 겸 CEO

## DevOps

DevOps는 사람, 프로세스, 기술을 한데 묶어주는 협업 기반의 접근 방식으로, 이를 통해 고품질 서비스 및 애플리케이션을 더 신속하게 제공할 수 있습니다.

## AI/ML에서 중요한 이유

조직에서는 AI/ML 모델은 물론 이를 사용하는 애플리케이션을 빠르고 효율적으로 개발하고 배포해야 합니다. 그러나 팀 간 협업의 부족으로 데이터 과학 프로젝트의 87% 정도가 프로덕션 단계까지 진행되지 못합니다.<sup>6</sup> DevOps 접근 방식은 AI/ML 팀, 애플리케이션 개발자, IT 운영 팀 사이의 협업을 촉진하여 ML 기반 애플리케이션에 대한 프로덕션 시간을 가속화합니다. 주로 **지속적 통합/지속적 제공(CI/CD)** 파이프라인 형태로 제공되는 자동화는 신속하고, 점진적이며, 반복적인 변화를 통해 애플리케이션 개발 라이프사이클을 더 빠르게 진행합니다.

## 모범 사례 및 권장 사항

DevOps는 기술만이 아닌, 사람과 프로세스 모두를 아우릅니다. DevOps 접근 방식을 AI/ML 라이프사이클을 전체에 적용해 보세요. 플랫폼과 툴에서 자동화는 물론 Argo, Tekton, Jenkins, Spinnaker와 같은 오픈소스 기술을 활용해 CI/CD 파이프라인을 구축합니다.



“이 솔루션이 아니었다면 적절한 수준의 분석과 효율성을 달성하는 데 그야말로 수백만 년이 필요했을 겁니다. Red Hat OpenShift는 DevOps 팀 전반에서 새로운 애플리케이션의 배포를 더욱 손쉽게 진행할 수 있도록 지원해줍니다.”<sup>7</sup>

Jochen Thaeder 박사  
DXC Technology, 고성능 데이터 기반 개발(Data-Driven Development, D3) 플랫폼 부문 수석 아키텍트

4 Red Hat, “[2020 Red Hat 글로벌 고객 기술 전망](#),” 2019년 11월

5 Red Hat 보도 자료 “[레드햇, 레드햇 오픈소스로 AI 및 머신러닝 워크플로우와 AI 기반 지능형 애플리케이션 가속화](#),” 2020년 3월 24일

6 VentureBeat, “[데이터 과학 프로젝트의 87%가 프로덕션까지 진행되지 못하는 이유\(Why do 87% of data science projects never make it into production?\)](#),” 2019년 7월 19일

7 Red Hat 고객 사례, “[데이터 플랫폼을 통해 자율 주행 분야에서 앞서 나가는 글로벌 기업](#),” 2020년 4월



# 하이브리드 클라우드 플랫폼 및 데이터 파이프라인

## 하이브리드 클라우드 플랫폼

하이브리드 클라우드 플랫폼은 온사이트, 엣지, 클라우드 환경에서 툴, 애플리케이션 및 모델을 개발, 배포, 관리할 수 있는 통합 소프트웨어 기반을 제공합니다.

## AI/ML에 중요한 이유

AI/ML 모델, 소프트웨어, 애플리케이션에는 개발 및 배포를 위한 인프라가 필요합니다. 일관된 하이브리드 클라우드 플랫폼은 인프라의 모든 부분에서 AI/ML 모델과 애플리케이션을 동일한 방식으로 개발, 테스트, 배포, 관리할 수 있도록 지원하여 유연성을 높여줍니다. 또한, IT 제어를 유지 관리하는 동시에 리소스 제공을 가속화하는 셀프서비스 기능도 제공합니다. 마지막으로, 일관된 플랫폼은 제3사 벤더, 오픈소스 커뮤니티, 사용할 수 있는 모든 커스텀 개발 툴의 기술 통합을 위한 기반이 되어줍니다.

## 모범 사례 및 권장 사항

하드웨어 가속화, 광범위한 AI/ML 및 애플리케이션 개발 툴, 에코시스템, 통합된 DevOps 및 운영 관리 기능을 지원하는 보안 중심 플랫폼을 선택하세요. 오픈소스 플랫폼을 선택하면 더 많은 통합 기회와 유연성을 확보할 수 있습니다.



"HCA Healthcare를 비롯한 의료 서비스 업계는 디지털 트랜스포메이션의 시작점에 있습니다. 우리가 구축하고 있는 플랫폼을 통해 실시간으로 모든 사이트에서 데이터를 수집하여 알고리즘을 구동하고 그 결과를 병상 치료에 적용할 수 있도록 도와줍니다."<sup>8</sup>

Edmund Jackson 박사  
HCA Healthcare, 최고 데이터 과학자

## 데이터 파이프라인

데이터 파이프라인은 AI/ML 모델 개발, 학습 및 추론을 위한 데이터 세트의 수집, 준비, 저장, 액세스 방법을 제공합니다.

## AI/ML에 중요한 이유

데이터는 모든 AI/ML 이니셔티브의 핵심 구성 요소로서 모델의 학습, 테스트, 운영을 위해 필요합니다. 그러나 Red Hat 고객의 22%는 AI/ML 이니셔티브에서 관련 데이터에 액세스하는 것이 가장 큰 과제라고 밝혔습니다.<sup>9</sup> 데이터 파이프라인은 개별적인 데이터 소스를 연결하고, 사용할 데이터를 준비해, AI/ML 엔지니어와 애플리케이션 개발자가 액세스할 수 있는 리포지토리에 이를 배치합니다. 또한 사용된 데이터를 임시 스토리지에서 영구 스토리지 아카이브로 이동할 수도 있습니다.

## 모범 사례 및 권장 사항

기존 데이터베이스, 데이터 레이크, 기타 리포지토리에 연결할 기술을 찾아보세요. 표준화된 애플리케이션 프로그래밍 인터페이스(API)와 대기 시간이 짧은 고대역폭 네트워킹은 AI/ML 라이프사이클 전반에서 데이터에 더 쉽게 액세스할 수 있도록 해줍니다. Apache Spark, Kafka, Presto와 같은 오픈소스 데이터 스트리밍, 조작, 분석 툴과 통합하면 데이터를 더욱 효율적으로 관리할 수 있습니다. 또한, 비즈니스를 보호하려면 데이터 거버넌스 기능과 통합 보안 기능을 제공하는 기술을 선택해야 합니다.



BMW Group은 쿠버네티스 기반 데이터 플랫폼을 사용해 약 230PB의 가용 스토리지에

액세스하고, 최대 2억 4000만 km의 테스트 데이터를 시뮬레이션합니다.<sup>10</sup>

<sup>8</sup> Red Hat 고객 사례, "Red Hat 소프트웨어를 사용해 예측 분석을 개발한 HCA Healthcare," 2019년 5월

<sup>9</sup> Red Hat, "2020 Red Hat 글로벌 고객 기술 전망," 2019년 11월

<sup>10</sup> Red Hat 고객 사례, "데이터 플랫폼을 통해 자율 주행 분야에서 앞서 나가는 글로벌 기업," 2020년 4월



# AI/ML을 위해 유연한 오픈 기반 구축

오픈소스 리더인 Red Hat은 완전한 기술 포트폴리오, 입증된 전문성, 전략적 파트너십을 제공하여 고객이 AI/ML 목표를 달성할 수 있도록 지원합니다. 광범위한 파트너 기술 에코시스템을 통해 프로덕션 레디 AI/ML 환경은 물론, 신속한 도입을 위한 서비스와 교육 기반을 제공합니다.

**Red Hat® OpenShift®**는 자동화된 운영을 통해 하이브리드 클라우드 배포를 관리하는 엔터프라이즈 레디 쿠버네티스 컨테이너 플랫폼입니다. 온디맨드 컴퓨팅 리소스, 하드웨어 가속 지원, 온사이트 및 클라우드 인프라 전반의 일관성은 팀의 성공에 필요한 속도와 유연성을 제공합니다. 셀프서비스 프로비저닝은 AI/ML 팀이 IT 팀의 개입 없이 리소스에 액세스할 수 있도록 지원합니다. NVIDIA 그래픽 처리 장치(GPU) 통합은 모델링 및 추론 태스크를 가속화합니다. 또한 협업 기능 덕분에 데이터 과학자는 컨테이너화된 모델링 결과를 생성하고 이를 동료 및 소프트웨어 개발자와 동일한 방식으로 공유할 수 있습니다. 빌트인 DevOps 기능은 AI/ML 기반의 지능형 애플리케이션 개발을 간소화합니다.

**Red Hat Middleware** 포트폴리오를 통해 통합된 애플리케이션 개발, 제공, 통합, 자동화 환경을 구축할 수 있습니다. 데이터 통합 서비스는 효과적인 데이터 파이프라인을 구축할 수 있도록 지원하고, 런타임 서비스는 애플리케이션 개발을 간소화해줍니다. 프로세스 자동화 툴과 서비스는 지능형 애플리케이션과 ML/DL 모델에 액세스하여 비즈니스 프로세스와 의사 결정을 간소화하고 자동화합니다.

**Red Hat Enterprise Linux®**, **Red Hat Virtualization**, **Red Hat OpenStack® Platform**, **Red Hat OpenShift Container Storage**, **Red Hat Ceph® Storage**를 비롯한 Red Hat 플랫폼 및 스토리지 제품은 하이브리드 운영을 위해 확장 가능한 소프트웨어 정의 인프라를 제공합니다.

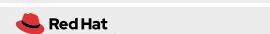
**Red Hat 인증 파트너 에코시스템**은 간편하고 자동화된 배포 및 라이프사이클 관리를 위해 선택한 AI/ML과 애플리케이션 개발 툴을 이러한 아키텍처에 통합하도록 지원합니다.

Red Hat은 **Kubeflow** 및 **Open Data Hub** 오픈소스 커뮤니티에 적극적으로 참여하고 있습니다. 이러한 커뮤니티에서는 Apache Spark, Kafka, TensorFlow, Jupyter 노트북, 기타 인기 데이터 및 ML 툴을 사용하여 청사진, 툴킷, 워크플로우를 제공합니다.



“Red Hat의 훌륭한 동료들과 협력하면서, 우리는 자연어 처리 및 머신 러닝과 같은 새로운 툴을 사용해 비정형 데이터에서 의료 업계를 뒤바꿀 새로운 인사이트를 발견할 수 있습니다.”

Jonathan Perlin 박사  
HCA Healthcare, 최고 의료 책임자

 ML/DL 및 DevOps 툴	
 ML/DL 데이터 파이프라인 및 소스	
 하이브리드 클라우드 플랫폼	
 하드웨어 가속화	
 인프라 엔드 포인트	    

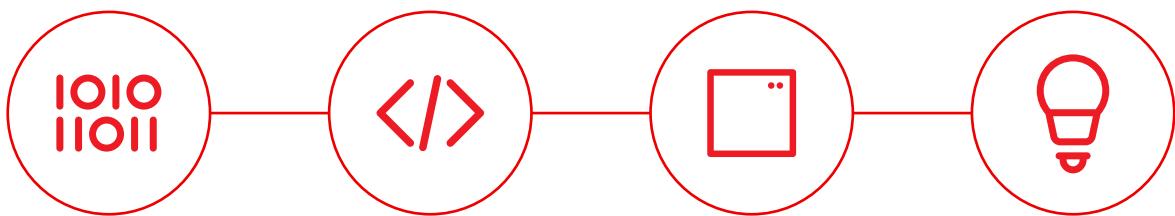


# 데이터에서 더 많은 가치를 창출할 준비가 되셨습니까?

AI, ML, DL은 비즈니스의 거의 모든 측면을 혁신적으로 변화시키고 있습니다. Red Hat은 지능형 애플리케이션의 개발 및 제공을 가속화할 프로덕션 레디 AI/ML 환경을 구축해 고객의 비즈니스 목표 달성을 지원합니다.

효과적인 AI/ML 배포로 얻을 수 있는 비즈니스 이점을 알아보세요.

[openshift.com/ai-ml](https://openshift.com/ai-ml)



## Red Hat Services로 빠르게 시작하세요.

Red Hat 전문가와 협력해 AI/ML 프로젝트를 신속하게 시작하세요. Red Hat은 조직에서 AI/ML을 더욱 신속하게 도입할 수 있도록 컨설팅과 교육 서비스를 제공합니다.

- AI/ML 서비스 자세히 알아보기: [red.ht/ai-consulting-services](https://red.ht/ai-consulting-services)
- 무료 디스커버리 세션을 예약해 보세요. <https://www.redhat.com/ko/services/consulting>