AI를 위한 스토리지:

# 수집에서 통찰에 이르는 빠른 방법

01
개요

02
IA(정보 아키텍처)가 없는 AI는 존재하지 않습니다

03
확실한 기반 위에 구축합니다

04
사례 연구: 경쟁 우위 확보

05
결론

## 개요

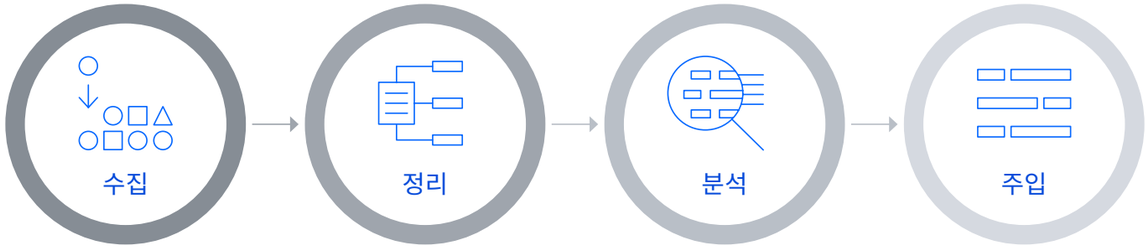
오늘날 조직은 AI(인공 지능)가 혁신과 생산성 향상에 이르는 빠른 방법이라는 점에 동의합니다. 대부분의 조직은 이미 AI의 모든 잠재력을 시험, 적용, 구현, 현실화하는 중이며, 이에 따라 향후 수년 동안 기업의 AI 솔루션 투자는 엄청나게 증가할 것으로 예상됩니다.

성공적인 모든 AI 프로젝트는 적합한 데이터 취득으로부터 다양한 AI 활용으로 이어지는 다단계 과정을 거쳐왔습니다.

지출 비용의 **40%** 2019년까지 디지털 변환은 AI 서비스를 사용할 것입니다[1]

**47억9천만 달러**

2019년 AI에 대한IT 스토리지[1]




데이터를 쉽게 사용할 수 있도록 만들기

신뢰성 있는 분석 기반 만들기

머신 러닝을 사용한 확장성 통찰

프로세스에 대한 예측, 자동화 및 확장성

AI를 적용하는 데 어려움이 전혀 없는 것은 아닙니다. 오픈 소스 그리고 상업용 개발자 도구와 프레임워크 덕분에 최초의 AI 프로젝트 또는 개념 증명 배포가 간단하게 되었습니다. 그러나 조직은 다음과 같은 AI 개발 팀 지원[1] 또는 AI 워크로드 제품화 배포와 확장 단계에서 해결해야 할 과제가 있습니다.

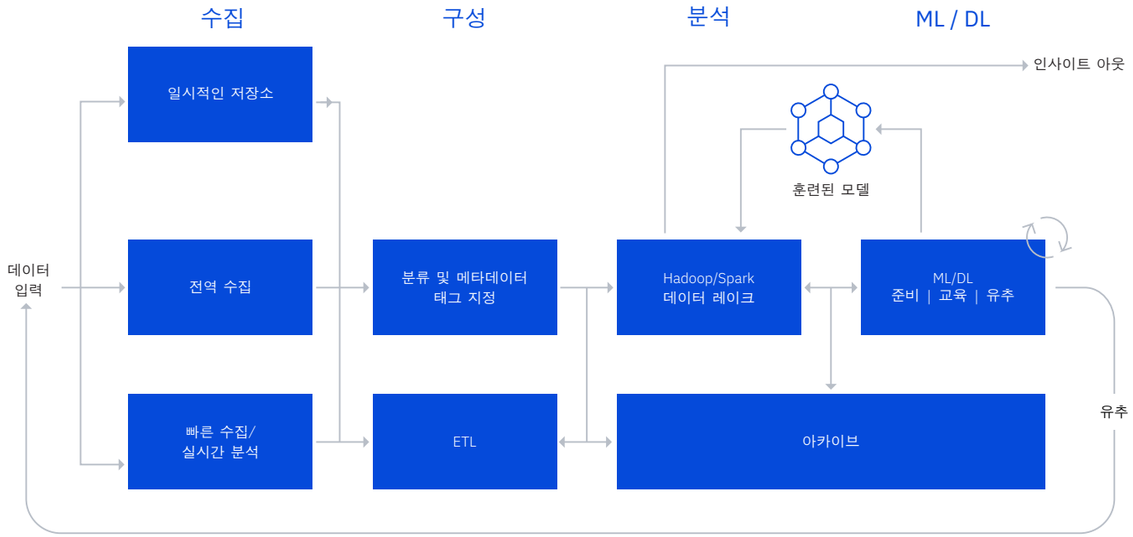
- **데이터 크기와 품질.** AI에는 고품질의 다양한 레이블이 구성된 데이터 입력이 필요합니다. 동적 데이터 특성을 지닌 여러 가지 데이터 소스에서 적합한 데이터 세트 식별 과정은 어려울 수 있습니다.
- **고급 데이터 관리.** AI 프로젝트에서 데이터 세트의 정리 및 추적 작업은 AI 모델의 정밀도를 개선하기 위해 데이터 세트를 반복적으로 시험, 재사용 및 확장해야 하는 개발자가 해결해야 할 과제입니다.
- **기술 공백.** AI 서비스의 수요가 증가함에 따라 숙련된 전문가에 대한 수요도 그에 따라 증가합니다. AI는 여전히 상대적으로 새로운 분야이므로 데이터 과학 생산성에 맞는 숙련된 기술자와 모범 사례를 찾기 어렵습니다.

수많은 조직이 어떻게 할지 확신하지 못하며, 이윤 창출을 위해 AI/ML을 가장 잘 활용하는 방법을 명확히 이해하지 못하는 것은 놀라운 일이 아닙니다. 바로 이런 경우를 위해 IBM®이 진행하는 모든 단계에서 도움을 드립니다.

데이터 크기와 품질, 첨단 데이터 관리, 기술 공백 등은 AI 개발팀을 지원하거나 AI 워크로드를 배포할 때 조직이 직면하는 중요한 도전 과제의 사례입니다.

# IA(정보 아키텍처)가 없는 AI는 존재하지 않습니다

AI 파이프라인(데이터를 수집, 정리 및 분석하며 궁극적으로, 데이터로부터 AI-주도 통찰력을 만들기 위해 모델을 학습시키는 방법)은 효율적인 데이터 과학의 핵심 요소입니다. 사용하는 AI 파이프라인의 효율성은 이 같은 도전 과제에 대해 적합한 IT 인프라를 사용하여 대처하는 것과 직접적으로 관련됩니다.




## 데이터 액세스 통합

데이터의 효율적인 사용에 있어서 특히, AI의 경우 데이터 창고는 큰 장애물입니다. AI 프로젝트에서 데이터 수집은 시간이 매우 오래 걸리는 단계가 될 수 있습니다. 모든 AI 프로젝트 전체에서 데이터 세트 정리 및 분류에 대한 기술 투자를 활용해야 합니다. 이때 중복을 최소화하고, 효율을 개선하며, 여러 개의 프로젝트에서 상호 공유하는 데이터를 만들고, 다양한 데이터 분석 도구를 지원하려면 데이터 및 스토리지 아키텍처가 필요합니다.

## 데이터 처리 성능

AI 모델의 정밀도는 좋은 데이터를 입력한 후 충분한 컴퓨팅 리소스로 분석한 결과물입니다. 대규모 데이터 세트를 빠르게 분석하기 위해 AI는 GPU(Graphics Processing Unit)를 종종 사용합니다. IT 인프라는 데이터를 처리하는 컴퓨팅 리소스의 용량에 잘 맞도록 스토리지 성능을 반드시 조정되어야 합니다. 이와 유사하게, 데이터 워크로드의 적절한 분산 특성이 필요한 경우, 실시간 통찰을 위해 스트리밍 데이터를 사용하기도 합니다.

## 컨테이너 지원을 통한 민첩성

AI 프로젝트는 일반적으로 가볍고, 빠르게 배포할 수 있으며, 여러 개의 프로그램과 스크립트를 조합할 수 있으므로 컨테이너를 사용하여 관리합니다. 초기 시도로부터 제품화 등급의 AI로 빠르게 확장하려면 Kubernetes와 Red Hat® OpenShift®와 함께 작동할 수 있는 영구 스토리지가 필요합니다. 컨테이너는 개발을 단순화할 뿐만 아니라 IT 인프라에 민첩성도 제공하므로 기업 AI 서비스의 수요 증가에 잘 맞습니다.

# 확실한 기반 위에 구축합니다

AI 작업 기준이 확장되면 복잡해질 것 같지만, 그렇지 않아도 됩니다. 확실한 기반 위에 구축되는 AI 프로젝트는 보다 손쉽고 더욱 성공하게 됩니다. IBM Storage for AI가 이와 같은 기반을 제공합니다. AI 워크로드 배포와 관련된 최상위 비즈니스 도전 과제에 대처하여 AI 생산에 적합한 빠른 방법을 제공하는 제품군을 갖추고 있습니다.

## IBM Spectrum 확장

IBM Spectrum® Scale은 귀사의 스토리지 인프라에 맞게 자동으로 확장 및 통합되는 고성능 파일 시스템 솔루션입니다. 본 제품은 최적의 스토리지 티어로 신속하게 그리고 효율적으로 파일 데이터를 이동하는 소프트웨어 방식을 통해 성능과 비용의 균형을 맞춥니다. IBM Spectrum Scale을 사용하면 AI 적용을 단순화하는 통합 데이터 기반을 이용하여 어디에서도 데이터에 액세스가 가능하므로 데이터를 안전하게 수집 및 정리할 수 있습니다.

IBM Spectrum Scale 살펴보기

## IBM Cloud Object Storage

IBM Cloud™ Object Storage는 클라우드 네이티브 응용프로그램 및 AI 프레임워크에 적합한 성능과 확장성을 제공합니다. 본 제품은 테라바이트로부터 엑사바이트까지의 용량 및 처리량을 손쉽게 확장하는 안전한 소프트웨어 방식의 스토리지 플랫폼입니다. IBM Cloud Object Storage는 최신 클라우드 개발 환경을 사용하며, 데이터 보안 또는 고성능 로컬 데이터도 필요한 팀에 이상적인 솔루션입니다.

IBM Cloud Object Storage 살펴보기

## IBM Spectrum Discover

IBM Spectrum Discover는 공용 클라우드를 포함한 여러 개의 스토리지 플랫폼 전체에서 메타데이터를 신속하게 수집, 통합, 색인 작성이 가능한 최신 메타데이터 관리 소프트웨어입니다. 본 제품을 사용하면 데이터 과학자와 스토리지 팀이 대량의 다양한 비정형 데이터로부터 메타데이터를 효율적으로 통합하고, 목록을 작성 및 다양화하여 통찰력을 확장시킬 수 있으므로 생산성이 상승합니다.

IBM Spectrum Discover 살펴보기

# 사례 연구: 경쟁 우위 확보

경쟁력을 엄청나게 강화시키는 데이터의 역량을 활용하십시오. AI는 데이터의 가치를 증명하는 최고의 핵심 요소이며, 다음과 같이 혁신적인 새로운 방법으로 비즈니스를 변환시켜 드립니다.

- 미래의 결과 예측 및 구체화
- 가치가 높은 작업에 역량 투입 최적화
- 의사 결정, 프로세스 및 경험의 자동화
- 비즈니스 모델 재구성

전체 데이터 수명 주기 관리를 개선하고, AI로 향하는 단계를 가속화하고, 조직을 변환하기 위해 IBM Storage를 사용한 일부 고객의 사례는 다음과 같습니다.

**결과:**

표준 유전체분석 파이프라인에서 런타임 96% 절감

동일한 작업을 대량으로 수행하는 상용 솔루션 사용 비용 1/3 절감

개념 설계로부터 클라우드에서 완벽하게 작동하는 IBM HPC 환경을 구축 하는 시간: 2주

사례 연구 보기

## L7 Informatics

**혁신적인 연구를 위한 고성능 Genomic Cloud 제공사입니다**

유전체학 - 유기체의 전체 DNA 집합을 연구하는 분야이며, 방대한 분량의 데이터 처리를 위해 과학자가 필요합니다. 따라서, 수많은 조직이 생성되는 방대한 분량의 데이터 처리를 위해 노력하고 있습니다.

L7 Informatics는 IBM과 함께 IBM Spectrum Storage 기술을 사용하는 제품을 사용하여 HPC(High Performance Computing) 환경을 구축하는 팀을 구성하여 다음과 같이 대처했습니다.

- 데이터 통합
- 방대한 분량의 비정형 데이터 처리
- 병목 현상이 없는 데이터의 병렬 액세스 처리
- 탄력적인 데이터 이동을 위한 티어링 기본 제공
- 분석 및 장기 보관을 위해 연구실로부터 클라우드를 향해 끊김없는 마이그레이션 제공

## Birmingham 대학교

**데이터 관리를 통해 혁신적인 연구 선도**

최신 연구 시뮬레이션에서는 그 어느 때보다 많은 데이터가 발생됩니다. 이와 같이 날로 증가하는 수요에 대처하기 위해 Birmingham 대학교는 IBM Spectrum Scale 및 IBM Spectrum Protect를 배포하여 다음과 같이 대처했습니다.

- 여러 개의 스토리지 시스템 전체에서 단일 데이터 관리 주체를 제공합니다
- 워크로드와 플랫폼을 조정할 때, 통제 불가능한 복잡성을 유발하지 않으면서 가격-성능 의사 결정을 할 수 있습니다
- 연구원이 적합한 곳에 애플리케이션을 즉각 배포할 수 있습니다

당사는 AI와 딥 러닝을 사용하는 기술을 적용 및 개발하는 분야를 포함한 매우 다양한 분야의 연구소를 지원합니다. 예를 들어, 당사는 Centre of Membrane Proteins and Receptors [비교] 프로젝트를 위해 Nottingham 대학교와 협력하고 있습니다. 최첨단 현미경에서 생성된 초고해상도 이미지를 분석하는 이 프로젝트는 보다 나은 심혈관 질환, 호흡기 질환 및 암 예방 및 치료법을 연구하고 있습니다.

Birmingham 대학교 연구 컴퓨팅 인프라 아키텍처 담당자, Simon Thompson

사례 연구 보기

# 결론

AI로 향하는 과정은 하나의 성공적인 개념 입증으로부터 시작하며, 조직 전체에서 빠르게 확장될 수 있습니다. 생산성과 적용성을 주도하는 독특한 데이터 요구사항에 최적화된 강력하고 민첩한 IT 기반을 만들면 이러한 과정 탐색을 성공적으로 시작할 수 있습니다. 적합한 스토리지 플랫폼은 AI 프로젝트에 필요한 성능, 확장성, 유연성을 반드시 제공해야 합니다. 이러한 기반을 구축하는 귀사의 의사 결정은 향후 모든 단계에 보다 광범위한 영향을 미치고 궁극적으로 귀사의 성공을 판가름합니다. 바로 이런 점 때문에 초기부터 적합한 파트너와 협력하는 것이 중요합니다.

IBM Storage for AI는 데이터 파이프라인의 종단간 최적화를 제공하여 데이터 거버넌스를 개선하고 통찰력에 이르는 시간을 단축합니다. 업계 최정상의 제품군, 혁신성과 검증된 리더십을 모두 갖춘 IBM이 인프라를 구축해 드립니다. 이를 통해 데이터를 관리하고, AI 워크로드를 조절하며, AI 역량을 사용하여 궁극적으로 더욱 깊은 통찰력을 주도하여 최선의 비즈니스 결과를 만드십시오.

IBM Storage for AI가 수집부터 통찰에 이르는 단계를 빠르게 하는 방법을 알려 드립니다.

IBM Storage for AI 안내

## 자원

1. 인지/AI 워크로드 예측을 위한 IDC Worldwide Storage, 2018-2022년

© 저작권 IBM Corporation 2019년. Direitos Restritos para Usuários do Governo dos EUA - Uso, duplicação ou divulgação restritos pelo documento GSA ADP Schedule Contract com a IBM Corp. 참고: IBM 웹 페이지에는 준수해야 할 다른 소유권 고시 및 저작권 정보가 들어 있을 수 있습니다.

IBM, IBM Spectrum, IBM Cloud, Red Hat, OpenShift, IBM 로고와 ibm.com은 전 세계 많은 관할지에 등록된 International Business Machines Corp.,의 상표입니다. 기타 제품 및 서비스 이름은 IBM 또는 기타 회사의 상표입니다. 현재 IBM 상표 목록은 웹(www.ibm.com/legal/copytrade.shtml)의 “저작권 및 상표 정보”를 참조하십시오.