



생성형 AI 워크로드 평가

AWS 권장 가이드



AWS 권장 가이드: 생성형 AI 워크로드 평가

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon의 상표 및 트레이드 드레스는 Amazon 외 제품 또는 서비스와 함께, Amazon 브랜드 이미지를 떨어뜨리거나 고객에게 혼동을 일으킬 수 있는 방식으로 사용할 수 없습니다. Amazon이 소유하지 않은 기타 모든 상표는 Amazon과 제휴 관계이거나 관련이 있거나 후원 관계와 관계없이 해당 소유자의 자산입니다.

Table of Contents

소개	1
이 가이드의 목적	1
대상 고객 및 이점	2
범위	2
목표 비즈니스 성과	4
평가 고려 사항 및 사전 조건	7
명확한 사용 사례로 시작	7
비즈니스 조정 보장	8
거버넌스 및 감독 구현	8
데이터 및 기술 사전 요구 사항 해결	8
컴퓨팅 리소스 요구 사항 고려	8
개인 정보 보호 및 보안 영향 해결	8
이해관계자를 조기에 참여시키기	9
반복 및 학습	9
생성형 AI 워크로드 평가 설문지	10
준비 상태	10
사용 사례	12
아키텍처	15
스토리지	15
규정 및 규정 준수	16
통합	17
테스트	19
배포 및 자동화	20
데이터 전략	22
평가 인사이트를 실행 가능한 결과로 변환	25
다음 단계	27
FAQ	28
기본 목표는 무엇입니까?	28
누가 이 평가를 사용해야 합니까?	28
주요 구성 요소는 무엇입니까?	28
이렇게 하면 아키텍처를 정의하는 데 어떻게 도움이 됩니까?	28
어떤 이점이 있나요?	28
이를 성공적으로 구현하려면 어떻게 해야 합니까?	29
어떤 문제가 있나요?	29

규제 및 규정 준수 요구 사항은 무엇입니까?	29
이해관계자의 역할은 무엇입니까?	29
성공을 어떻게 측정할 수 있습니까?	29
조직 크기에 따라 접근 방식이 어떻게릅니까?	29
리소스	31
문서 기록	32
용어집	33
#	33
A	34
B	36
C	38
D	41
E	45
F	47
G	48
H	49
I	51
L	53
M	54
O	58
P	60
Q	63
R	63
S	66
T	69
U	71
V	71
W	72
Z	73
.....	lxxiv

생성형 AI 워크로드 평가

Tabby Ward 및 Deepak Dixit, Amazon Web Services(AWS)

2024년 11월([문서 기록](#))

생성형 AI 워크로드 평가는 생성형 AI 워크로드를 생성하거나 업데이트하기 위한 조직의 준비 상태를 평가하고 개선하는 것을 목표로 하는 전략적 방법입니다. 생성형 AI를 비즈니스 운영에 통합하면 사물의 작동 방식이 크게 변화하고 새로운 효율성과 기능을 제공할 수 있으므로 이 평가가 중요합니다. 그러나 생성형 AI를 성공적으로 채택하려면 현재 시스템을 철저히 이해하고 미래를 위한 명확한 계획을 세우는 것이 중요합니다.

생성형 AI 워크로드는 텍스트, 이미지, 코드 또는 기타 데이터 유형과 같은 새 콘텐츠를 생성할 수 있는 인공지능 모델을 사용하는 컴퓨팅 작업을 말합니다. 이러한 워크로드에는 일반적으로 상당한 컴퓨팅 성능, GPUs와 같은 특수 하드웨어, 훈련 및 추론을 위한 대규모 데이터 세트가 필요합니다. 생성형 AI 워크로드를 작업에 통합하면 다음과 같은 몇 가지 문제가 발생합니다.

- 인프라 요구 사항: 생성형 AI 모델에 필요한 중요한 컴퓨팅 리소스 및 특수 하드웨어 프로비저닝.
- 데이터 관리: 대규모 데이터 세트를 처리하면서 데이터 품질, 개인 정보 보호 및 규정 준수를 보장합니다.
- 기술 격차: AI 기술 및 모델 배포에 대한 전문 지식이 부족합니다.
- 윤리적 고려 사항: AI 생성 콘텐츠의 편향, 공정성 및 투명성 해결.
- 통합 복잡성: 생성형 AI를 기존 워크플로 및 레거시 시스템에 원활하게 통합합니다.
- 비용 관리: 잠재적 이점과 높은 구현 및 운영 비용의 균형을 맞춥니다.

이러한 문제를 해결하려면 신중한 계획, 인프라 및 인재에 대한 투자, 구현에 대한 전략적 접근 방식이 필요합니다.

이 가이드의 목적

생성형 AI는 많은 산업에서 빠르게 중요한 구성 요소가 되고 있습니다. 이는 혁신적인 기회를 제공하지만 통합, 규정 준수 및 확장성 측면에서도 문제를 야기합니다. 많은 조직이 취약한 기술적 기반, 변화에 대한 저항, 데이터 품질 문제로 인해 AI를 완전히 활용하기 위해 노력하고 있습니다. 생성형 AI 워크로드 평가는 현대화 요구 사항을 식별하고, 구현 범위를 정의하고, 레거시 시스템과 사고에 도전하여 이러한 문제를 해결합니다. 또한 최소 실행 가능 제품(MVPs) 결정하는 데 도움이 되며 대상 솔루션 아키텍처를 개발하여 AI 채택에 대한 구조화되고 전략적인 접근 방식을 보장하는 데 도움이 됩니다.

이 가이드는 조직이 생성형 AI 기술을 채택하는 복잡성을 탐색하는 데 도움이 되는 구조화된 접근 방식의 역할을 합니다. 이 가이드는 처음부터 요구 사항을 명확하게 정의하는 대신 다음을 지원합니다.

- 조직 내 생성형 AI의 잠재적 사용 사례 식별.
- 생성형 AI 채택에 대한 조직의 준비 상태 평가.
- 사용 사례 목표 및 확장 목표를 정의하고 구체화합니다.
- 생성형 AI 구현의 범위 및 요구 사항 결정.
- 대상 솔루션 아키텍처 개발.

대상 고객 및 이점

이 평가는 생성형 AI 워크로드 현대화의 기술적 측면을 평가하려는 솔루션 아키텍트, 엔터프라이즈 아키텍트 및 애플리케이션 아키텍트를 위해 특별히 설계되었습니다. 또한 팀의 전반적인 준비 상태, 리소스 할당 및 활성화 요구 사항을 측정하려는 프로그램 및 인력 관리자에게도 유용합니다. 업계 모범 사례는 AI 채택 준비 상태를 보장하기 위한 포괄적인 평가의 중요성을 강조합니다. 여기에는 아키텍처, 스토리지, 규정 준수, 통합, 테스트, 배포 및 자동화 평가가 포함됩니다.

범위

다음 주제는 생성형 AI 워크로드 평가 방법의 범위 내에 있습니다.

- 현재 생성형 AI 기술 및 모델(예: 대규모 언어 모델, 이미지 생성 모델)
- 생성 기법을 사용하는 좁은 AI 애플리케이션
- 생성형 AI를 기존 시스템 및 워크플로와 통합
- 생성형 AI 모델 훈련 및 미세 조정을 위한 데이터 전략
- 현재 생성형 AI 애플리케이션에 대한 윤리적 고려 사항 및 책임 있는 AI 사례
- 프로덕션 환경에서 생성형 AI를 위한 테스트 및 배포 전략
- 생성형 AI 구현을 위한 보안 및 개인 정보 보호 고려 사항
- 생성형 AI 워크로드의 성능 최적화 및 확장성
- 다양한 산업에서 생성형 AI의 사용 사례 및 애플리케이션
- 생성형 AI 출력 및 품질 보증 프로세스 평가

다음 주제는 범위를 벗어납니다.

- 인공 일반 지능(AGI) 및 인공 지능(ASI) 시나리오
- 현재 생성형 모델을 넘어 AI의 투기적 미래 발전
- AI의 양자 컴퓨팅 애플리케이션
- Neuromorphic 컴퓨팅 및 brain-computer 인터페이스
- AI 시스템의 인식 및 자기 인식
- 현재 생성형 AI 애플리케이션을 넘어서는 고급 AI의 장기적인 사회적 영향
- 가상의 미래 AI 기술을 위한 규제 프레임워크
- 기계에서 지능과 의식의 특성에 대한 철학적 논쟁
- AI의 극단적인 옻지 사례 또는 고도로 투기적인 사용 사례
- 독점 AI 모델 또는 아키텍처의 세부 기술 사양

목표 비즈니스 성과

생성형 AI 워크로드 평가는 생성형 AI 워크로드를 성공적으로 현대화하는 데 중요한 여러 대상 결과를 제공하는 것을 목표로 합니다. 이러한 결과를 통해 조직은 AI 기술을 효과적이고 효율적으로 통합할 준비가 잘 되어 있습니다.

각 대상 결과에 대해 생성형 AI 워크로드 평가는 다음에 중점을 둡니다.

- 상호 종속성: 현대화 프로세스의 결과와 다른 측면 간의 상호 종속성을 식별하고 명확히 합니다. 여기에는 현대화에 대한 전체적인 접근 방식을 보장하기 위해 한 결과가 다른 결과에 어떤 영향을 미칠 수 있는지 이해하는 것이 포함됩니다.
- 이해관계자 조정: 다양한 이해관계자를 각 결과에 맞추기 위한 전략을 간략하게 설명합니다. 여기에는 각 결과의 가치와 영향을 다양한 조직 수준과 부서에 전달하여 동의와 지원을 촉진하는 것이 포함됩니다.
- 우선순위 지정: 여러 사용 사례 또는 결과가 식별되는 경우 비즈니스 영향, 리소스 요구 사항 및 전략적 조정과 같은 요소를 기반으로 우선순위를 지정하기 위한 프레임워크를 제공합니다.
- 지속적인 개선: 각 결과에 대해 지속적인 평가 및 개선을 위한 메커니즘을 설정합니다. 이를 통해 현대화 노력이 적응력을 유지하고 변화하는 기술 환경과 비즈니스 요구 사항에 대응할 수 있습니다.

다음은 각 대상 결과에 대한 자세한 설명입니다.

대상 아키텍처

- 정의: 평가는 생성형 AI 워크로드를 위한 명확하고 확장 가능한 대상 아키텍처를 정의하는 데 도움이 됩니다.
- 구성 요소: 여기에는 적절한 클라우드 서비스 선택, 데이터 파이프라인 설계, 시스템 상호 운용성 보장이 포함됩니다.
- 이점: 잘 정의된 아키텍처는 확장성, 안정성 및 성능 최적화를 지원하고 현대화를 위한 강력한 기반을 제공합니다.

고객 준비 상태

- 평가: 조직의 인프라, 프로세스 및 문화의 현재 상태를 평가하여 생성형 AI 현대화 채택에 대한 준비 상태를 결정합니다.
- 기준: 여기에는 기술 역량, 데이터 품질 및 변화를 수용하려는 조직의 의지 평가가 포함됩니다.

- 결과: 격차와 개선 영역을 식별하면 조직이 최신 솔루션과 기술로 원활하게 전환할 준비가 된 것입니다.

사용 사례 목표 및 확장 목표

- 사용 사례 목표는 특정 비즈니스 문제 또는 기회에 초점을 맞춰 대상 솔루션 구현을 위한 명확한 목표를 설정합니다.

생성형 AI 현대화의 맥락에서 사용 사례 목표는 조직이 생성형 AI 솔루션을 구현하여 달성하고자 하는 구체적이고 측정 가능한 목표를 말합니다. 이러한 목표는 일반적으로 더 광범위한 비즈니스 목표에 부합하며 조직 내 특정 과제 또는 기회를 해결하는 데 중점을 둡니다. 사용 사례 목표의 예는 다음과 같습니다.

- 생성형 AI 기반 챗봇을 사용하여 고객 서비스 응답 시간을 50% 단축합니다.
- 생성형 AI 지원 코드 분석을 통해 코드 검토 효율성을 30% 개선합니다.
- 생성형 AI 패턴 인식을 사용하여 사기 탐지 정확도를 25% 높입니다.
- 확장 목표는 생성형 AI 현대화가 조직 내에서 달성할 수 있는 것의 경계를 넓히는 야심찬 목표를 정의합니다.
- 영향: 달성 가능한 목표와 야심찬 목표를 모두 설정하면 생성형 AI 현대화 이니셔티브를 전략적 비즈니스 목표에 맞추고 혁신을 장려하는 데 도움이 됩니다.

노력 추정

- 목적: 정확한 노력 추정은 리소스 계획을 지원하고 프로젝트가 예산 내에서 적시에 제공되도록 합니다.
- 범위: 생성형 AI 현대화 계획을 구현하는 데 필요한 리소스, 시간 및 예산을 추정합니다.
- 요인: 기술적 복잡성, 통합 문제 및 잠재적 위험을 고려합니다.

활성화 요구 사항

- 훈련 및 개발: 성공적인 생성형 AI 현대화 채택에 필요한 기술과 지식을 식별합니다.
- 리소스: 훈련 프로그램, 워크숍 및 기타 지원 활동의 필요성을 결정합니다.
- 결과: 직원이 필요한 기술을 갖추도록 하면 생성형 AI 현대화 이니셔티브의 효율성이 향상되고 장기적인 성공을 지원할 수 있습니다.

구현 계획

- 로드맵: 생성형 AI 현대화를 달성하는 데 필요한 단계를 설명하는 세부 계획을 개발합니다.
- 마일스톤: 진행 상황을 추적하기 위한 주요 마일스톤 및 결과물을 정의합니다.
- 이점: 명확한 구현 계획은 방향과 책임을 제공하고 생성형 AI 현대화에 대한 구조화된 접근 방식을 용이하게 합니다.

평가 고려 사항 및 사전 조건

명확한 사용 사례로 시작

생성형 AI가 해결할 수 있는 특정 비즈니스 문제 또는 기회를 식별합니다. 전략적 비즈니스 목표에 부합하는 사용 사례에 집중하고 측정 가능한 이점을 제공합니다. 솔루션 아키텍처가 여러 시나리오의 패턴 역할을 할 수 있도록 조직 내에서 일반적으로 직면하는 문제를 대상으로 하는 사용 사례의 우선순위를 정합니다.

잠재적 생성형 AI 애플리케이션을 전반적으로 이해하여 평가 프로세스를 시작하는 것은 유용하지만 필수는 아닙니다. 이 가이드에 포함된 [설문](#)은 사용 사례가 잘 정의된 조직부터 광범위한 아이디어만 있는 조직에 이르기까지 다양한 수준의 준비 상태를 수용합니다. 평가 프로세스는 다음을 수행합니다.

- 이러한 초기 사용 사례 아이디어를 구체화하고 명확히 합니다.
- 새로운 잠재적 사용 사례를 식별합니다.
- 각 사용 사례에 대해 구체적이고 측정 가능한 목표를 개발합니다.
- 각 사용 사례의 실행 가능성과 잠재적 영향을 평가합니다.

가상의 예를 살펴보겠습니다. 금융 서비스 회사가 생성형 AI 현대화를 탐색하기로 결정합니다. 먼저 고객 서비스 및 사기 탐지 프로세스를 개선하는 광범위한 아이디어로 시작합니다.

- 초기 평가: 설문은 생성형 AI 채택을 위한 현재 시스템, 데이터 품질 및 조직 준비 상태를 평가하는 데 도움이 됩니다.
- 사용 사례 세분화: 평가 프로세스를 통해 초기 아이디어를 다음과 같은 두 가지 특정 사용 사례로 세분화합니다.
 - 고객 문의를 위한 생성형 AI 기반 챗봇 구현
 - 실시간 트랜잭션 사기 탐지에 생성형 AI 사용
- 목표 설정: 각 사용 사례에 대해 특정 목표를 정의합니다.
 - 6개월 이내에 고객 서비스 응답 시간 40% 단축
 - 사기 탐지 정확도 20% 향상 및 오탐률 15% 감소
- 확장 목표: 또한 다음과 같은 야심 찬 목표를 설정합니다.
 - AI 지원 응답으로 고객 만족도 80% 달성
 - 새로운 사기 패턴을 식별하는 예측 사기 탐지 모델 개발

- MVP 정의: 이 설문은 즉각적인 가치를 제공하는 필수 기능에 중점을 두고 각 사용 사례에 대한 MVP를 결정하는 데 도움이 됩니다.
- 대상 아키텍처: 마지막으로 하나 또는 두 사용 사례를 모두 지원하는 대상 아키텍처를 개발하고 기존 시스템과의 확장성 및 통합을 보장합니다.

비즈니스 조정 보장

생성형 AI 이니셔티브를 전반적인 비즈니스 전략 및 목표에 맞게 조정합니다. 각 사용 사례에 대해 생성형 AI가 비즈니스 성장, 효율성 또는 혁신에 어떻게 기여하는지 보여주는 명확한 가치 제안을 개발합니다. 생성형 AI 구현이 핵심 성과 지표(KPIs)에 미치는 영향을 측정하는 지표를 설정합니다.

거버넌스 및 감독 구현

다기능 조향 위원회를 만들어 생성형 AI 이니셔티브를 감독합니다. 윤리적 고려 사항과 잠재적 편향을 해결하여 책임감 있는 AI 사용에 대한 정책과 지침을 개발합니다. 생성형 AI 프로젝트에 대한 검토 프로세스를 수립하여 조직 표준 및 규제 요구 사항을 준수하는지 확인합니다.

데이터 및 기술 사전 요구 사항 해결

데이터 품질을 평가 및 개선하고 데이터 거버넌스 관행을 구현하여 생성형 AI 모델에 대한 안정적인 입력을 보장합니다. 생성형 AI 요구 사항에 맞는 데이터 수집, 스토리지 및 관리를 다루는 데이터 전략을 개발합니다. 생성형 AI 워크로드에 필요한 데이터의 볼륨과 속도를 지원하도록 데이터 인프라를 평가하고 개선합니다.

컴퓨팅 리소스 요구 사항 고려

현재 IT 인프라를 평가하고 생성형 AI 워크로드의 컴퓨팅 용량 격차를 식별합니다. 클라우드 서비스 또는 온프레미스 고성능 컴퓨팅 클러스터와 같은 옵션을 고려하여 확장 가능한 컴퓨팅 리소스를 계획합니다. 리소스 할당을 최적화하여 훈련 워크로드와 추론 워크로드의 성능과 비용 효율성의 균형을 맞춥니다.

개인 정보 보호 및 보안 영향 해결

강력한 보안 조치를 구현하여 생성형 AI 훈련 및 운영에 사용되는 민감한 데이터를 보호합니다. 개인 정보를 처리할 때 일반 데이터 보호 규정(GDPR) 또는 캘리포니아 소비자 개인정보 보호법(CCPA)과

같은 데이터 보호 규정을 준수해야 합니다. 생성형 AI 기능의 무단 액세스 또는 오용을 방지하기 위해 안전한 모델 배포 및 모니터링을 위한 프로토콜을 개발합니다.

이해관계자를 조기에 참여시키기

리더십 동의와 지원을 얻기 위해 처음부터 주요 이해관계자를 참여시킵니다. 특히 생성형 AI 워크로드에 대한 현대화 이니셔티브의 이점과 잠재적 영향을 명확하게 전달합니다. 이해관계자가 생성형 AI 기술과 그 영향을 이해하는 데 도움이 되는 교육 및 리소스를 제공합니다.

반복 및 학습

대상 솔루션을 구체화할 수 있는 증분 접근 방식을 채택합니다. 피드백 루프를 사용하여 워크로드 아키텍처 및 프로세스를 지속적으로 개선합니다. 생성형 AI 구현의 성능과 영향을 정기적으로 평가하고 실제 결과와 변화하는 비즈니스 요구 사항에 따라 필요에 따라 전략을 조정합니다.

생성형 AI 워크로드 평가 설문지

다음 섹션에서는 조직의 생성형 AI 워크로드 현대화의 다양한 측면을 평가하는 데 사용할 수 있는 질문을 제공합니다. 이 포괄적인 설문은 사용 사례, 아키텍처, 스토리지, 규정 준수, 통합, 테스트, 배포 및 데이터 전략을 비롯한 주요 영역에서 질문을 통해 생성형 AI 워크로드를 채택하고 구현할 수 있는 조직의 준비 상태를 평가합니다. 이 설문지는 기술 인프라에서 규제 고려 사항에 이르기까지 생성형 AI 구현의 중요한 측면을 해결하여 AI 현대화 여정의 강점, 격차 및 기회를 식별하는 데 도움이 됩니다.

섹션:

- [준비 상태](#)
- [사용 사례](#)
- [아키텍처](#)
- [스토리지](#)
- [규정 및 규정 준수](#)
- [통합](#)
- [테스트](#)
- [배포 및 자동화](#)
- [데이터 전략](#)

Microsoft Excel 형식으로 설문을 다운로드하여 정보를 기록하는 데 사용할 수도 있습니다.

 [설문 다운로드](#)

준비 상태

질문	응답의 예
이러한 워크로드에 활용할 수 있는 AWS 계정이 있습니까?	예 또는 아니요.
와 기존 엔터프라이즈 계약이 있습니까 AWS?	예 또는 아니요.
생성형 AI 워크로드를 처리하기 위한 현재 클라우드 인프라의 확장성은 어느 정도입니까?	클라우드 인프라는 대규모 생성형 AI 워크로드를 효율적으로 처리하도록 설계된 컴퓨팅 리소

질문	응답의 예
	스 및 분산 스토리지 시스템에 대한 자동 조정 기능을 통해 확장성이 뛰어납니다.
사전 처리 및 대규모 특성 엔지니어링을 위한 데이터 파이프라인 기능이 있습니까?	데이터 파이프라인은 대규모 데이터 사전 처리 및 기능 엔지니어링을 위해 Apache Spark와 같은 분산 처리 프레임워크를 사용하며 배치 및 스트리밍 데이터 처리를 모두 지원합니다.
계정 프로비저닝 및 관리 기능이 있습니까?	예 또는 아니요.
생성형 AI 기술을 채택하기 위한 조직의 AI 리더십과 준비 상태를 어떻게 설명하시겠습니까?	우리 조직은 AI 교육 프로그램에 많은 투자를 했으며 대부분의 기술 인력은 기본 AI/ML 교육을 완료했습니다. 조직은 생성형 AI를 비롯한 새로운 기술을 수용하는 혁신 문화를 가지고 있습니다.
조직 내에는 어떤 AI/ML 전문 지식이 있으며 어떻게 배포되나요?	숙련된 데이터 과학자 및 ML 엔지니어가 있는 전용 AI Center of Excellence가 있습니다. AI 리더레이션이 되고 생성형 AI 사용 사례를 식별하도록 다양한 사업부의 도메인 전문가의 기술을 높입니다.
클라우드 프로그램 목표, 이점 및 비용을 설명하는 개괄적인 비즈니스 사례가 있습니까?	예 또는 아니요.
솔루션을 프로덕션 환경에 적용하기 위한 일정은 어떻게 됩니까?	주, 월 등.
주요 이해관계자(예: CFO, CIT/CTO, COO)가 자금 지원을 약속했습니까?	예 또는 아니요.
생성형 AI 이니셔티브에서 데이터 보호 규정을 준수하려면 어떻게 해야 하나요?	AI 팀과 긴밀히 협력하는 전담 규정 준수 팀이 있습니다. 당사는 정기적으로 개인 정보 영향 평가를 수행하고, 설계 원칙에 따라 데이터 보호를 구현하고, 모든 생성형 AI 프로젝트에 대한 세부 데이터 처리 레코드를 유지합니다.

질문	응답의 예
<p>새로운 생성형 AI 기술과 통합되는 기존 시스템은 얼마나 성숙합니까?</p>	<p>IT 아키텍처는 새로운 생성형 AI 기술을 유연하게 통합할 수 있는 마이크로서비스 및 APIs를 기반으로 합니다. 이러한 시스템은 상호 운용성을 보장하기 위해 일반적인 데이터 형식 및 프로토콜을 기반으로 표준화됩니다.</p>
<p>ML 모델을 운영할 때 어떤 경험이 있으며 생성형 AI 시스템에 어떻게 적용될 수 있습니까?</p>	<p>자동화된 모델 배포 파이프라인, 모니터링 시스템 및 A/B 테스트 프레임워크를 비롯한 MLOps 관행을 수립했습니다. 이러한 관행은 대규모 생성형 AI 모델의 고유한 요구 사항을 처리하도록 조정되고 있습니다.</p>

사용 사례

질문	응답의 예
<p>사용 사례의 주요 목표 또는 성공 기준은 무엇입니까?</p>	<p>고객 지원 응답 시간을 개선하려면 판매 전환을 늘리고 제품 권장 사항을 개선하세요. 또한 사용자 만족도, 작업 완료율, 응답 품질 등을 개선합니다.</p>
<p>이 사용 사례는 조직의 전략적 목표에 어떻게 부합하나요?</p>	<p>이는 고객 서비스의 응답 시간을 줄여 고객 만족도를 높이는 전략적 목표에 부합합니다.</p>
<p>사용 사례에 대한 예상 데이터 또는 요청량은 얼마입니까?</p>	<p>초당 500개의 트랜잭션(TPS).</p>
<p>생성형 AI 워크로드를 지원하려면 어떤 유형의 데이터 소스가 필요합니까?</p>	<p>내부 구조화 데이터베이스(고객 레코드, 판매 데이터 등), 문서, 이메일 및 소셜 미디어의 구조화되지 않은 텍스트 데이터, 음성 및 이미지 인식 작업을 위한 오디오 및 비디오 파일, IoT 디바이스 및 센서의 실시간 스트리밍 데이터, 보강을 위한 퍼블릭 데이터 세트 및 APIs.</p>

질문	응답의 예
이러한 소스에서 데이터를 얼마나 자주 업데이트하거나 새로 고쳐야 합니까?	트랜잭션 데이터베이스: 거의 실시간 업데이트, 문서 리포지토리: 일일 배치 업데이트, 소셜 미디어 피드: 시간별 업데이트, IoT 센서 데이터: 연속 실시간 스트리밍, 퍼블릭 데이터 세트: 월별 또는 분기별 업데이트.
생성형 AI 모델에 입력으로 필요한 데이터 형식은 무엇입니까?	구조화된 데이터: CSV, JSON 및 SQL 데이터베이스 테이블, 텍스트 데이터: 일반 텍스트, PDF 및 HTML, 이미지 데이터: JPEG, PNG 및 TIFF, 오디오 데이터: WAV 및 MP3, 비디오 데이터: MP4 및 AVI.
생성형 AI 워크로드의 주요 데이터 품질 문제는 무엇입니까?	완전성: 중요한 필드가 누락되지 않았는지 확인, 정확성: 데이터 정확성 확인 및 오류 제거, 일관성: 소스 간에 균일한 형식 및 값 유지, 적시성: 실시간 추론을 위해 데이터가 최신 상태인지 확인, 관련성: 데이터가 특정 생성형 AI 작업과 일치하는지 확인.
주요 성능 요구 사항(예: 응답 시간, 처리량, 정확도)은 무엇입니까?	95% 정확도, < 500ms 응답 시간, 초당 1,000 개의 요청을 처리할 수 있는 기능. 높은 정확도 (95% 이상), 중간 정확도(80~90%), 최선의 노력 등.
이 사용 사례의 성공을 측정하는 다른 KPIs가 있습니까?	주요 KPIs에는 오류율 감소, 트랜잭션당 시간 절감, 고객 만족도 점수가 포함됩니다.
얼마나 많은 모델 정확도가 필요하고 비용과 어떻게 균형을 이루어야 합니까?	중간 비용의 높은 정확도(>90%), 낮은 비용의 중간 정확도(70~80%) 등입니다.
생성형 AI 솔루션의 주요 사용 사례 또는 시나리오에는 무엇입니까?	고객 서비스 챗봇, 콘텐츠 생성, 제품 권장 사항 등.
생성형 AI 시스템의 대상 사용자 또는 페르소나는 무엇입니까?	고객 서비스 에이전트, 마케팅 팀, 직원, 최종 사용자 등.
요청 또는 사용자의 예상 볼륨은 얼마입니까?	일일 요청 1,000건, 월별 활성 사용자 10,000명.

질문	응답의 예
특정 사용 사례 제약 또는 요구 사항이 있나요?	실시간 응답, 다국어 지원, 데이터 개인 정보 보호 등.
생성형 AI 솔루션을 개발하고 유지 관리하기 위한 예산이 할당되어 있습니까?	초기 개발 비용은 200,000 USD로 추정되며 연간 유지 관리 비용은 50,000 USD입니다.
이 사용 사례의 예상 투자 수익률(ROI) 및 회수 기간은 어떻게 됩니까?	3년 동안 예상 ROI는 150%이고 페이백 기간은 18개월입니다.
고려해야 할 숨겨진 비용이나 잠재적 절감이 있나요?	잠재적 절감에는 초과 근무 비용 절감이 포함됩니다. 숨겨진 비용에는 직원을 위한 추가 교육이 포함될 수 있습니다.
이 생성형 AI 솔루션의 확장성과 향후 확장 가능성은 무엇입니까?	이 솔루션은 향후 다른 부서로 확장할 수 있는 운영에 따라 확장되도록 설계되었습니다.
생성형 AI 모델에서 공정성을 보장하고 편향을 완화하려면 어떻게 해야 하나요?	다양한 데이터 수집, 정기적인 편향 감사, 편향 완화 기법 구현을 통해 편향을 완화할 계획입니다.
윤리적 문제나 의도하지 않은 결과를 해결하기 위해 어떤 프로세스를 마련하고 있습니까?	당사는 확립된 AI 인시던트 대응 계획, 정기적인 윤리적 위험 평가, 직원을 위한 익명 보고 시스템, 외부 윤리적 전문가와의 협업, 피드백을 기반으로 배포된 모델의 지속적인 모니터링 및 조정을 통해 윤리적 우려를 관리할 것입니다.
조직의 다양한 프로젝트와 부서에서 생성형 AI 워크로드 평가의 우선순위를 정하고 순서를 정하는 데 어떻게 접근하나요?	모든 부서에서 개괄적인 설문 조사를 수행하여 잠재적인 생성형 AI 사용 사례를 식별하고 비즈니스 영향, 기술적 타당성 및 윤리적 고려 사항의 세 가지 주요 기준에 따라 평가합니다. 잠재적 영향이 높고, 기술적 장벽이 낮으며, 윤리적 문제가 최소화된 프로젝트에는 우선 순위가 부여됩니다.

아키텍처

질문	응답의 예
어떤 유형의 생성형 AI 모델 또는 아키텍처가 고려되고 있습니까?	변환기, 컨볼루션 신경망(CNN), 반복 신경망(RNN), 의사 결정 트리 등.
데이터 및 계산의 예상 규모 또는 볼륨은 얼마입니까?	수백만 명의 사용자, 페타바이트의 데이터 등.
훈련 및 추론을 위한 하드웨어 요구 사항(예: CPUs 또는 GPUs)은 무엇입니까?	하이엔GPU, CPU 클러스터, 클라우드 인스턴스 등.
생성형 AI 모델은 시간이 지남에 따라 어떻게 업데이트되거나 재학습되나요?	지속적 학습, 주기적 재학습, 수동 업데이트 등을 통해.
데이터 사전 처리 및 기능 엔지니어링 요구 사항은 무엇입니까?	텍스트 정리, 이미지 보강, 기능 선택 등.
생성형 AI 시스템은 엣지 케이스, 이상치 또는 신뢰도가 낮은 입력을 어떻게 처리하나요?	인적 감독으로의 대체, 요청 설명 등을 통해.
생성형 AI 애플리케이션의 지연 시간 요구 사항은 무엇입니까?	실시간, 거의 실시간, 배치 처리 등.

스토리지

질문	응답의 예
훈련 데이터는 어디에 저장되나요?	클라우드 스토리지(예: Amazon S3, 파일 스토리지, 블록 스토리지 또는 객체 스토리지), 온프레미스 스토리지 등.
훈련 데이터 및 모델 아티팩트(예: 용량, 내구성, 가용성)의 스토리지 요구 사항은 무엇입니까?	페타바이트 규모의 스토리지, 높은 내구성(99.999999999% 내구성),고가용성 등.

질문	응답의 예
훈련 데이터 및 모델 아티팩트에 대한 데이터 보존 및 백업 요구 사항은 무엇입니까?	x년, 일별 백업, 오프사이트 백업 등의 데이터 보존.
AI 훈련 데이터 세트를 저장하는 데 주로 사용되는 파일 형식(예: CSV, JSON, Parquet, HDF5)은 무엇입니까?	정형 데이터용 Parquet 파일과 이미지 및 텍스트와 같은 대형 다차원 배열 및 비정형 데이터용 HDF5. TFRecord와 같은 특수 형식을 사용하여 훈련 중에 데이터 로드를 최적화합니다.
개별 파일, 데이터베이스 또는 특수 AI 데이터 형식 사용 등 훈련 데이터 세트는 어떻게 구성되어 있습니까?	중소 규모 데이터 세트는 유연성을 위해 객체 스토리지에 개별 Parquet 파일로 저장됩니다. 대규모 데이터 세트는 분산 데이터베이스(Cassandra)에 저장되어 규모를 처리합니다.
생성형 AI 훈련 데이터에 특히 데이터 압축 또는 인코딩 기술을 사용하십니까?	테이블 형식 데이터의 경우 Parquet에서 사용할 수 있는 사전 인코딩 및 비트 패키징 기술을 사용합니다. 이미지의 경우 모델에 최적화된 품질 설정과 함께 손실된 JPEG 압축을 사용합니다.
다양한 교육 데이터 세트 반복의 버전 관리 및 저장을 어떻게 처리하나요? 이는 전체 스토리지 요구 사항에 어떤 영향을 미칩니까?	ML 플랫폼과 통합된 데이터 버전 관리 시스템(DVC)을 사용합니다.

규정 및 규정 준수

질문	응답의 예
생성형 AI 솔루션(예: GDPR, HIPAA, PCI-DSS)에 대한 관련 규정 또는 규정 준수 요구 사항은 무엇입니까?	개인 데이터 처리를 위한 GDPR, 의료 데이터에 대한 HIPAA, 결제 데이터에 대한 PCI-DSS 등.
조직에서 채택한 윤리적 생성형 AI 지침 또는 프레임워크는 무엇입니까?	책임감 있는 AI 지침을 구현했습니다. 모든 생성형 AI 프로젝트는 승인 및 배포 전에 윤리적 검토를 거칩니다.

질문	응답의 예
생성형 AI 시스템의 보안 요구 사항은 무엇입니까?	데이터 암호화, 보안 네트워크 통신, 정기적인 보안 감사.
데이터 프라이버시 및 보호에 대한 요구 사항은 무엇입니까?	데이터 익명화, 암호화, 액세스 제어 등.
솔루션이 민감한 데이터 또는 기밀 데이터를 처리하기 위한 요구 사항은 무엇입니까?	엄격한 액세스 제어, 데이터 마스킹, 데이터 레지던시 요구 사항 등.
사용자 인증 및 권한 부여는 어떻게 처리되나요?	OAuth, API 키, SSO(Single Sign-On) 및 RBAC(역할 기반 액세스 제어)를 사용합니다.
프로덕션 환경에서 솔루션을 어떻게 모니터링하고 관리하나요?	Prometheus 및 Datadog과 같은 모니터링 도구를 사용하여 ELK 스택과 같은 로깅 도구, 알림 시스템 등을 사용합니다.

통합

질문	응답의 예
생성형 AI 솔루션을 기존 시스템 또는 데이터 소스와 통합하기 위한 요구 사항은 무엇입니까?	REST APIs, 메시지 대기열, 데이터베이스 커넥터 등.
생성형 AI 솔루션을 위해 데이터를 어떻게 수집하고 사전 처리하나요?	배치 처리, 스트리밍 데이터, 데이터 변환 및 기능 엔지니어링을 사용합니다.
생성형 AI 솔루션의 출력은 어떻게 소비되거나 다운스트림 시스템과 통합되나요?	API 엔드포인트, 메시지 대기열, 데이터베이스 업데이트 등을 통해
생성형 AI 솔루션에 사용할 수 있는 이벤트 기반 통합 패턴은 무엇입니까?	메시지 대기열(예: Amazon SQS , Apache Kafka, RabbitMQ), pub/sub 시스템, 웹후크, 이벤트 스트리밍 플랫폼.
생성형 AI 솔루션을 다른 시스템과 연결하는 데 사용할 수 있는 API 기반 통합 접근 방식은 무엇입니까?	RESTful APIs, GraphQL APIs, SOAP APIs(레거시 시스템용).

질문	응답의 예
<p>생성형 AI 솔루션 통합에 사용할 수 있는 마이크로서비스 아키텍처 구성 요소는 무엇입니까?</p>	<p>서비스 간 통신, API 게이트웨이, 컨테이너 오케스트레이션(예: Kubernetes)을 위한 서비스 메시입니다.</p>
<p>생성형 AI 솔루션을 위해 하이브리드 통합을 구현하려면 어떻게 해야 하나요?</p>	<p>실시간 업데이트를 위한 이벤트 기반 패턴, 기록 데이터를 위한 배치 처리, 외부 시스템 통합을 위한 APIs 결합하여</p>
<p>생성형 AI 솔루션 출력을 다운스트림 시스템과 통합하려면 어떻게 해야 하나요?</p>	<p>API 엔드포인트, 메시지 대기열, 데이터베이스 업데이트, 웹훅 및 파일 내보내기를 통해</p>
<p>생성형 AI 솔루션을 통합하려면 어떤 보안 조치를 고려해야 합니까?</p>	<p>인증 메커니즘(예: OAuth 또는 JWT), 암호화(전송 중 및 저장 중), API 속도 제한 및 액세스 제어 목록(ACLs).</p>
<p>LlamaIndex 또는 LangChain과 같은 오픈 소스 프레임워크를 기존 데이터 파이프라인 및 생성형 AI 워크플로에 통합하려면 어떻게 해야 하나요?</p>	<p>LangChain을 사용하여 특히 에이전트 및 메모리 관리 기능을 위한 복잡한 생성형 AI 애플리케이션을 구축할 계획입니다. 향후 6개월 이내에 생성형 AI 프로젝트의 60%가 LangChain을 사용하는 것을 목표로 합니다.</p>
<p>선택한 오픈 소스 프레임워크와 기존 데이터 인프라 간의 호환성을 보장하려면 어떻게 해야 하나요?</p>	<p>원활한 호환성을 보장하기 위해 전용 통합 팀을 만들고 있습니다. 3분기까지 목표는 현재 데이터 레이크 구조 내에서 효율적인 데이터 인덱싱 및 검색을 위해 LlamaIndex를 사용하는 완전 통합 파이프라인을 확보하는 것입니다.</p>
<p>신속한 프로토타입 제작 및 실험을 위해 LangChain과 같은 프레임워크의 모듈식 구성 요소를 어떻게 활용할 계획입니까?</p>	<p>개발자가 LangChain의 구성 요소를 사용하여 빠르게 프로토타입을 만들 수 있는 샌드박스 환경을 설정하고 있습니다.</p>
<p>빠르게 진화하는 이러한 오픈 소스 프레임워크의 업데이트와 새로운 기능을 따라잡기 위한 전략은 무엇입니까?</p>	<p>LangChain 및 LlamaIndex LlamaIndex에 대한 GitHub 리포지토리 및 커뮤니티 포럼을 모니터링 팀을 지정했습니다. 성능 개선 및 새로운 기능에 중점을 두고 분기별로 주요 업데이트를 평가하고 통합할 계획입니다.</p>

테스트

질문	응답의 예
테스트 요구 사항은 무엇입니까(예: 단위 테스트, 통합 테스트, end-to-end 테스트)?	개별 구성 요소에 대한 단위 테스트, 외부 시스템과의 통합 테스트, 중요한 시나리오에 대한 end-to-end 테스트 등.
생성형 AI 훈련을 위해 다양한 소스에서 데이터 품질과 일관성을 보장하려면 어떻게 해야 하나요?	자동화된 데이터 프로파일링 도구, 정기적인 데이터 감사 및 중앙 집중식 데이터 카탈로그를 통해 데이터 품질을 유지합니다. 소스 간 일관성을 보장하고 데이터 계보를 유지하기 위해 데이터 거버넌스 정책을 구현했습니다.
생성형 AI 모델은 어떻게 평가되고 검증되나요?	홀드아웃 데이터 세트, 인적 평가, A/B 테스트 등을 사용합니다.
생성형 AI 모델의 성능과 정확도를 평가하는 기준은 무엇입니까?	정밀도, 재현율, F1 점수, 복잡성, 인적 평가 등.
엣지 케이스와 모서리 케이스는 어떻게 식별되고 처리되나요?	포괄적인 테스트 제품군, 인적 평가, 적대적 테스트 등을 사용합니다.
생성형 AI 모델의 잠재적 편향은 어떻게 테스트 하나요?	인구통계학적 평등 분석, 기회 균등 테스트, 적대적 편향 제거 기법, 반사실적 테스트 등을 사용합니다.
모델 출력의 공정성을 측정하는 데 사용되는 지표는 무엇입니까?	서로 다른 영향 비율, 균등화된 확률, 인구통계학적 평등, 개별 공정성 지표 등입니다.
편향 감지를 위해 테스트 데이터 세트에서 다양한 표현을 보장하려면 어떻게 해야 합니까?	인구 집단 전반에서 계층화된 샘플링을 사용하고, 다양성 전문가와 협업하고, 합성 데이터를 사용하여 격차를 줄이는 등의 작업을 수행합니다.
배포 후 모델 공정성을 지속적으로 모니터링하기 위해 구현할 프로세스는 무엇입니까?	정기적인 공정성 감사, 자동화된 편향 감지 시스템, 사용자 피드백 분석, 업데이트된 데이터 세트를 사용한 주기적 재훈련 등.

질문	응답의 예
생성형 AI 모델의 교차 편향을 어떻게 해결하시겠습니까?	교차 공정성 분석, 하위 그룹 테스트, 교차성에 대한 도메인 전문가와의 협업 등을 사용합니다.
다양한 언어와 문화적 맥락에서 모델의 성능을 테스트하려면 어떻게 해야 하나요?	다국어 테스트 세트, 문화 전문가와의 협업, 현지화된 공정성 지표, 다문화 비교 연구 등을 사용합니다.

배포 및 자동화

질문	응답의 예
크기 조정 및 로드 밸런싱의 요구 사항은 무엇입니까?	지능형 요청 라우팅, 자동 조정 시스템, 모델 캐싱, 지연 로딩, 분산 스토리지 시스템과 같은 기술을 사용하여 빠른 콜드 스타트를 위한 최적화, 급증하고 예측할 수 없는 트래픽 패턴을 처리하도록 시스템 설계.
새 버전을 업데이트하고 출시하기 위한 요구 사항은 무엇입니까?	블루/그린 배포, 카나리 릴리스, 롤링 업데이트 등.
재해 복구 및 비즈니스 연속성을 위한 요구 사항은 무엇입니까?	백업 및 복원 절차, 장애 조치 메커니즘,고가용성 구성 등.
생성형 AI 모델의 훈련, 배포 및 관리를 자동화하기 위한 요구 사항은 무엇입니까?	자동 훈련 파이프라인, 지속적 배포, 자동 조정 등.
새로운 데이터를 사용할 수 있게 되면 생성형 AI 모델은 어떻게 업데이트되고 재학습되나요?	주기적 재훈련, 증분 학습, 전이 학습 등을 통해.
모니터링 및 관리를 자동화하기 위한 요구 사항은 무엇입니까?	자동 알림, 자동 조정, 자가 복구 등.
생성형 AI 워크로드에 선호하는 배포 환경은 무엇입니까?	AWS를 모델 훈련에 사용하고 온프레미스 인프라를 추론에 사용하여 데이터 레지던시 요구 사항을 충족하는 하이브리드 접근 방식입니다.

질문	응답의 예
생성형 AI 배포에 선호하는 특정 클라우드 플랫폼이 있나요?	AWS 서비스, 특히 모델 개발 및 배포를 위한 Amazon SageMaker AI와 파운데이션 모델을 위한 Amazon Bedrock.
생성형 AI 워크로드에 어떤 컨테이너화 기술을 고려하고 있나요?	하이브리드 환경 전체에서 이식성과 확장성을 보장하기 위해 Kubernetes로 오케스트레이션된 Docker 컨테이너를 표준화하려고 합니다.
생성형 AI 파이프라인에 CI/CD에 대해 선호하는 도구가 있습니까?	버전 관리 및 CI/CD 파이프라인용 GitLab은 자동화된 테스트 및 배포를 위해 Jenkins와 통합됩니다.
생성형 AI 워크플로를 관리하기 위해 어떤 오케스트레이션 도구를 고려하고 있나요?	워크플로 오케스트레이션, 특히 데이터 사전 처리 및 모델 훈련 파이프라인을 위한 Apache Airflow.
생성형 AI 워크로드를 지원하기 위한 온프레미스 인프라에 대한 특정 요구 사항이 있습니까?	온프레미스 추론 워크로드를 지원하기 위해 GPU 가속 서버 및 고속 네트워킹에 투자하고 있습니다.
다양한 환경에서 모델 버전 관리 및 배포를 어떻게 관리할 계획입니까?	MLflow를 모델 추적 및 버전 관리에 사용하고 이를 Kubernetes 인프라와 통합하여 환경에 원활하게 배포할 계획입니다.
생성형 AI 배포를 위해 어떤 모니터링 및 관찰성 도구를 고려하고 있나요?	지표 수집용 Prometheus와 시각화용 Grafana, 모델별 모니터링을 위한 추가 사용자 지정 로깅 솔루션.
하이브리드 배포 모델에서 데이터 이동 및 동기화를 어떻게 처리하나요?	온프레미스 스토리지와 간에 효율적인 데이터 전송을 AWS DataSync 위해를 사용하며 AWS, 훈련 주기에 따라 예약된 자동 동기화 작업을 사용합니다.
다양한 환경에서 생성형 AI 배포를 위해 어떤 보안 조치를 구현하고 있나요?	온프레미스 Active Directory와 통합된 클라우드 리소스에 IAM을 사용하여 end-to-end 암호화 및 네트워크 세분화를 구현하여 데이터 흐름을 보호합니다.

데이터 전략

질문	응답의 예
생성형 AI 워크로드에 중요한 특정 데이터 유형은 무엇이며, 이 중 현재 액세스할 수 있는 데이터 유형은 몇 퍼센트입니까?	고객 통화 로그 및 제품 리뷰 데이터는 매우 중요합니다. 현재 이러한 데이터 유형의 85%는 생성형 AI 프로젝트에 액세스할 수 있습니다.
데이터의 품질을 어떻게 보장하고 측정하나요?	완전성, 정확성, 일관성 및 적시성을 포함한 데이터 품질 지표를 구현했습니다. 자동화된 도구를 사용하여 이러한 지표를 정기적으로 평가하고 데이터 정리 및 보강을 위한 전담 팀이 있습니다.
생성형 AI 사용에 대한 품질 표준을 충족하는 데이터의 비율은 얼마나 됩니까?	현재 데이터의 78%가 품질 표준을 충족합니다. 데이터 정리 프로세스를 개선하여 향후 12개월 이내에 95%를 목표로 하고 있습니다.
이해관계자 간에 생성형 AI의 데이터 사용에 대한 신뢰를 구축하려면 어떻게 해야 하나요?	AI 위원회를 구현하고, AI 결정에 대한 명확한 설명을 제공하고, 투명성과 공정성을 보장하기 위해 분기별 AI 감사를 수행하고 있습니다.
데이터 소스 및 계보에 대한 설명서는 얼마나 포괄적입니까?	오리진, 업데이트 빈도 및 사용량을 포함하여 모든 데이터 소스에 대한 메타데이터가 포함된 세부 데이터 카탈로그를 유지합니다. 데이터 계보 도구를 사용하여 시스템 전반의 데이터 흐름 및 변환 방식을 추적합니다.
AI 모델의 편향을 방지하기 위해 데이터 세트의 다양성을 보장하려면 어떻게 해야 하나요?	다양한 인구 집단에서 데이터를 적극적으로 수집하고 데이터 세트에 표현 편향이 있는지 정기적으로 감사합니다. 또한 합성 데이터 생성 기술을 사용하여 과소 대표된 범주의 균형을 맞춥니다.
중요한 생성형 AI 모델의 데이터 새로 고침 빈도는 얼마이며 이 빈도를 어떻게 결정하나요?	중요 모델은 매주 새로 고쳐집니다. 이 빈도는 A/B 테스트 성능 지표에 따라 결정되며 새로 고침 간 성능 저하를 2% 이하로 줄이는 것을 목표로 합니다.

질문	응답의 예
얼마나 많은 버전의 중요 데이터 세트를 유지 관리하며 얼마나 오래 유지 관리하나요?	각 중요한 데이터 세트의 마지막 5개 버전을 유지하며, 각 버전에 대해 보존 기간은 18개월입니다.
생성형 AI 이니셔티브에 참여하고 데이터에 액세스할 수 있는 부서 간 팀은 몇 명입니까?	3개의 부서 간 팀이 있습니다. 각 팀에는 데이터 과학자, 도메인 전문가, 민속학자 및 비즈니스 분석가가 포함됩니다.
어떤 데이터 거버넌스 정책 및 관행이 마련되어 있습니까?	데이터 정책을 감독하는 부서 간 데이터 거버넌스 위원회가 있습니다. 거버넌스 프레임워크 준수를 보장하기 위해 역할 기반 액세스 제어, 데이터 분류 체계 및 정기 감사를 구현했습니다.
데이터 프라이버시를 보장하고, 적절한 동의를 얻고, 기밀성을 유지하기 위해 어떤 조치가 마련되어 있습니까?	GDPR 및 CCPA와 일치하는 포괄적인 데이터 프라이버시 프레임워크를 구현했습니다. 여기에는 데이터 사용에 대한 명시적 동의 획득, 데이터 익명화 기법 구현, 정기적인 개인 정보 영향 평가가 포함됩니다.
지난 분기에 편향에 대해 감사를 받은 AI 훈련 데이터 세트의 비율은 얼마나 됩니까?	지난 분기에 AI 훈련 데이터 세트의 70%가 편향에 대해 감사를 받았습니다. 100% 분기별 감사를 달성하기 위해 자동화된 편향 감지 도구를 구현하고 있습니다.
현재 데이터 처리 용량은 얼마이며 향후 생성형 AI 워크로드에 얼마나 필요할 것으로 예상하십니까?	현재 용량은 10TB/일입니다. 1년 내에 30TB/일이 필요할 것으로 예상되며이 수요에 맞게 인프라를 확장하고 있습니다.
생성형 AI 모델의 데이터 요구 사항과 데이터 프라이버시의 균형을 맞추기 위한 전략은 무엇입니까?	고급 익명화 기법과 합성 데이터 생성을 구현하고 있습니다. 목표는 내년에 개인 정보 보호 위험을 60% 줄이면서 AI에 사용할 수 있는 데이터를 40% 늘리는 것입니다.

질문	응답의 예
기계 학습(ML) 데이터 세트의 몇 퍼센트에 정확하게 레이블이 지정되어 있으며 목표 정확도는 얼마입니까?	현재 ML 데이터 세트의 85%에 정확하게 레이블이 지정되어 있습니다. 인적 레이블 지정 기법과 자동화된 레이블 지정 기법을 모두 사용하여 다음 분기 내에 95%의 정확도를 목표로 합니다.

평가 인사이트를 실행 가능한 결과로 변환

이 섹션에서는 설문 응답을 분석하고 이러한 인사이트를 사용하여 생성형 AI 현대화 이니셔티브의 대상 아키텍처 및 기타 주요 결과물을 형성하는 프레임워크를 제공합니다. 이 프레임워크는 데이터 수집과 구현 간의 격차를 해소하고 평가가 현대화 전략을 직접 알리고 주도하도록 합니다.

대상 아키텍처 정의:

- 설문 응답을 사용하여 클라우드 서비스 선택 및 데이터 파이프라인 설계에 대한 정보를 제공합니다.
- 가이드에 강조 표시된 대로 아키텍처 설계가 확장성과 상호 운용성을 지원하는지 확인합니다.

고객 준비 상태 평가:

- 현재 인프라, 프로세스 및 조직 문화와 관련된 설문 응답을 분석합니다.
- 격차를 식별하고 이를 해결하기 위한 계획을 수립합니다. MVP 성공에 중요한 격차의 우선순위를 정합니다.

사용 사례 및 확장 목표:

- 설문 응답에서 특정 비즈니스 문제를 추출하여 명확한 사용 사례 목표를 정의합니다.
- 생성형 AI 현대화에 대한 조직의 장기 비전에 부합하는 확장 목표를 설정합니다.

노력 추정:

- 설문 데이터를 사용하여 MVP와 전체 구현 모두에 대한 리소스, 시간 및 예산을 추정합니다.
- MVP로 시작하는 단계별 접근 방식을 만들고 후속 단계를 간략하게 설명합니다.

활성화 요구 사항:

- 설문 응답을 기반으로 기술 격차와 교육 요구 사항을 식별합니다.
- 즉각적인 MVP 요구 사항과 장기 생성형 AI 채택을 모두 지원하는 훈련 계획을 개발합니다.

구현 계획:

- MVP로 시작하여 전체 생성형 AI 현대화를 위한 단계를 간략하게 설명하는 포괄적인 로드맵을 생성합니다.

- 구현의 각 단계에 대한 명확한 마일스톤과 결과물을 정의합니다.

실제 단계:

- 우선순위 매트릭스: 설문 응답을 [6가지 결과에](#) 매핑하여 특성과 작업의 우선순위를 지정하는 데 도움이 되는 매트릭스를 생성합니다.
- 반복적 접근 방식: MVP를 계획된 일련의 릴리스에서 첫 번째 반복으로 설계합니다. 여기서 각 릴리스는 전체 대상 아키텍처를 향해 구축됩니다.
- 이해관계자 조정: 설문 결과를 사용하여 MVP 범위에 대한 이해관계자와 모든 결과를 달성하기 위한 단계별 접근 방식을 조정합니다.
- 지속적인 피드백 루프: MVP 배포 후 피드백을 수집하는 메커니즘을 구현하고 인사이트를 사용하여 후속 단계에 대한 계획을 구체화합니다.
- 애자일 구현: MVP에서 가장 중요한 성과부터 시작하여 시간 경과에 따른 모든 결과를 유연하게 처리할 수 있는 애자일 방법론을 채택합니다.

다음 단계

생성형 AI 워크로드 평가를 완료한 후 다음 단계를 따릅니다.

1. 세부 대상 아키텍처 제공

- 목표: 솔루션 아키텍트는 조직의 목표와 평가 결과에 맞는 포괄적인 대상 아키텍처를 생성합니다.
- 구성 요소: 이 아키텍처에는 확장성, 안정성 및 성능 최적화를 보장하기 위한 데이터 수집, 통합 지점 및 시스템 상호 운용성 설계가 포함되어 있습니다.

2. 사용 사례에 얼마나 구체적인 AWS 서비스 지 설명

- 서비스 매핑: 식별된 사용 사례에 AWS 서비스 가장 적합한 특정을 식별하고 매핑합니다.
- 이점: 이러한 서비스가 특정 비즈니스 요구 사항을 해결하고, 효율성을 높이고, 확장성을 제공하는 방법을 강조합니다.

3. 장단점이 포함된 선택적 대체 솔루션 제공

- 대안: 조직의 요구 사항을 충족할 수도 있는 대안 솔루션을 제시합니다.
- 분석: 비용, 복잡성, 비즈니스 목표 조정 등의 요소를 고려하여 각 대안의 장단점을 자세히 분석합니다.

4. 의 세부 가격 추정 제공 AWS 서비스

- 비용 분석: 잠재적 사용 시나리오 및 요금 모델을 AWS 서비스포함하여 제안된에 대한 자세한 비용 추정을 제공합니다.
- 예산 조정: 비용이 조직의 예산 제약 조건에 부합하고 재무적 영향에 대한 명확한 이해를 제공하는지 확인합니다.

5. 제안된 아키텍처에 대한 피드백 받기

- 이해관계자 참여: 이해관계자와 협력하여 제안된 아키텍처를 제시하고 피드백을 수집합니다.
- 반복적 개선: 피드백을 사용하여 솔루션을 개선 및 개선하고 모든 이해관계자의 요구와 기대치를 충족하는지 확인합니다.

FAQ

생성형 AI 워크로드 평가의 주요 목표는 무엇인가요?

평가의 주요 목표는 생성형 AI 워크로드를 현대화하고, 사용 사례를 식별하고, 대상 솔루션 아키텍처를 개발하기 위한 조직의 준비 상태를 평가하는 것입니다. 현대화 요구 사항을 정의하고, 구현 범위를 결정하고, 성공적인 생성형 AI 현대화를 준비하는 것을 목표로 합니다.

누가이 평가를 사용해야 합니까?

이 평가는 생성형 AI 현대화의 기술적 측면을 평가하려는 솔루션 아키텍트, 엔터프라이즈 아키텍트 및 애플리케이션 아키텍트를 위한 것입니다. 또한 프로그램 관리자와 인력 관리자가 전반적인 준비 상태, 리소스 할당 및 활성화 요구 사항을 측정하는 데도 유용합니다.

평가에서 평가되는 주요 구성 요소는 무엇입니까?

이 평가에서는 전반적인 준비 상태, 사용 사례, 아키텍처, 스토리지, 규정 및 규정 준수, 통합, 테스트, 배포 자동화 및 데이터 전략을 다룹니다. 이러한 구성 요소는 생성형 AI 현대화 채택을 위한 기술적 및 조직적 준비 상태를 결정하는 데 매우 중요합니다.

평가는 대상 아키텍처를 정의하는 데 어떻게 도움이 됩니까?

평가는 현재 시스템을 평가하고 개선 사항을 식별하기 위한 구조화된 접근 방식을 제공합니다. 이를 통해 비즈니스 목표 및 사용 사례 요구 사항에 맞는 적절한 기술을 선택하고 확장 가능한 아키텍처를 설계할 수 있습니다.

생성형 AI 워크로드 평가를 수행하면 어떤 이점이 있나요?

이점으로는 효율성 향상, 의사 결정 개선, 규정 준수 보장, 혁신 촉진 및 확장성 준비 등이 있습니다. 평가는 생성형 AI 현대화에 대한 전략적 접근 방식을 설정하고 위험을 완화하면서 잠재적 이점을 극대화합니다.

조직은 평가 후 성공적인 구현을 어떻게 보장할 수 있습니까?

조직은 정의된 마일스톤을 포함하는 명확한 구현 계획을 개발하고, 이해관계자를 조기에 참여시키고, 반복적인 접근 방식을 채택해야 합니다. CoE(Center of Excellence)를 설정하고 인재 개발에 집중하는 것도 모범 사례입니다.

조직에서 평가 중에 직면할 수 있는 문제는 무엇입니까?

문제에는 변화에 대한 저항, 데이터 품질 문제 및 규정 준수 복잡성이 포함될 수 있습니다. 이러한 문제를 해결하려면 혁신 문화를 조성하고, 데이터 준비를 보장하고, 강력한 보안 조치를 구현해야 합니다.

평가는 규제 및 규정 준수 요구 사항을 어떻게 해결합니까?

평가는 현재 규정 준수 측정값을 평가하고 격차를 식별합니다. 이를 통해 대상 솔루션이 관련 규정 및 데이터 개인 정보 보호법을 준수하고 보안 모범 사례를 통합하여 민감한 정보를 보호할 수 있습니다.

평가 프로세스에서 이해관계자 참여는 어떤 역할을 합니까?

이해관계자 참여는 동의를 얻고, 현대화 이니셔티브를 비즈니스 목표에 맞추고, 성공적인 구현을 보장하는 데 매우 중요합니다. 조기에 개입하고 이점을 명확하게 전달하는 것은 저항을 극복하고 지원을 장려하는 데 중요합니다.

조직은 평가 후 생성형 AI 현대화 이니셔티브의 성공을 어떻게 측정할 수 있습니까?

비즈니스 목표에 맞는 핵심 성과 지표(KPIs)를 사용하여 성공을 측정할 수 있습니다. 이러한 지표를 정기적으로 모니터링하고 평가하면 의사 결정을 안내하고 이해관계자에게 생성형 AI 현대화의 가치를 입증하는 데 도움이 됩니다.

다양한 규모(소규모, 중형 또는 엔터프라이즈) 또는 산업의 조직에서 평가 접근 방식은 어떻게 다릅니까?

소규모 조직:

- 포괄적인 평가를 위한 리소스와 전문 지식이 제한적일 수 있음

- 전사적 채택 대신 특정 고위험 사용 사례에 집중할 가능성이 높음
- 평가를 위해 타사 도구 및 서비스에 더 많이 의존할 수 있음
- 평가 프로세스는 덜 공식적이고 더 민첩할 수 있습니다.

중간 규모의 조직:

- 전용 IT 또는 데이터 팀이 있는 경우가 많지만 전문 AI 전문 지식이 부족할 수 있음
- 주요 부서의 파일럿 프로젝트부터 시작하여 단계별 접근 방식을 취할 수 있습니다.
- 혁신과 기존 시스템 및 프로세스의 균형을 맞춰야 함
- 평가에는 부서 간 팀이 포함될 수 있음

엔터프라이즈 조직:

- 일반적으로 포괄적인 평가를 위한 전용 AI/ML 팀과 더 많은 리소스 보유
- 기존 엔터프라이즈 시스템과의 복잡한 통합을 고려해야 함
- 고려해야 할 산업별 규제 요구 사항이 있을 수 있습니다.
- 평가는 종종 공식 거버넌스 프로세스를 포함합니다.

리소스

- [의 생성형 AI AWS](#)
- [AWS 는 AI 전략을 계획하기 위한 새로운 인공지능, 기계 학습 및 생성형 AI 가이드를](#) 제공합니다 (AWS 블로그 게시물).
- [에서 생성형 AI 애플리케이션을 빌드하는 모범 사례 AWS](#)(AWS 블로그 게시물)
- [의 생성형 AI 애플리케이션 빌더 AWS](#)(AWS 솔루션 라이브러리)
- [생성형 AI 기능](#)(AWS 보안 참조 아키텍처)
- [AWS 생성형 AI 모범 사례 프레임워크](#)(AWS Audit Manager 사용 설명서)
- [생성형 AI 서비스 선택](#)(AWS 결정 가이드)
- [Amazon Bedrock이란 무엇입니까?](#) (Amazon Bedrock 사용 설명서)
- [Amazon SageMaker AI란 무엇인가요?](#) (Amazon SageMaker AI 개발자 안내서)

문서 기록

아래 표에 이 가이드의 주요 변경 사항이 설명되어 있습니다. 향후 업데이트에 대한 알림을 받으려면 [RSS 피드](#)를 구독하십시오.

변경 사항	설명	날짜
최초 게시	—	2024년 11월 6일

AWS 권장 가이드 용어집

다음은 AWS 권장 가이드에서 제공하는 전략, 가이드 및 패턴에서 일반적으로 사용되는 용어입니다. 용어집 항목을 제안하려면 용어집 끝에 있는 피드백 제공 링크를 사용하십시오.

숫자

7가지 전략

애플리케이션을 클라우드로 이전하기 위한 7가지 일반적인 마이그레이션 전략 이러한 전략은 Gartner가 2011년에 파악한 5가지 전략을 기반으로 하며 다음으로 구성됩니다.

- 리팩터링/리아키텍트 - 클라우드 네이티브 기능을 최대한 활용하여 애플리케이션을 이동하고 해당 아키텍처를 수정함으로써 민첩성, 성능 및 확장성을 개선합니다. 여기에는 일반적으로 운영 체제와 데이터베이스 이식이 포함됩니다. 예: 온프레미스 Oracle 데이터베이스를 Amazon Aurora PostgreSQL 호환 에디션으로 마이그레이션합니다.
- 리플랫폼(리프트 앤드 리세이프) - 애플리케이션을 클라우드로 이동하고 일정 수준의 최적화를 도입하여 클라우드 기능을 활용합니다. 예: 온프레미스 Oracle 데이터베이스를 AWS 클라우드의 Amazon Relational Database Service(Amazon RDS) for Oracle로 마이그레이션합니다.
- 재구매(드롭 앤드 슝) - 일반적으로 기존 라이선스에서 SaaS 모델로 전환하여 다른 제품으로 전환합니다. 예: 고객 관계 관리(CRM) 시스템을 Salesforce.com으로 마이그레이션합니다.
- 리호스팅(리프트 앤드 시프트) - 애플리케이션을 변경하지 않고 클라우드로 이동하여 클라우드 기능을 활용합니다. 예: 온프레미스 Oracle 데이터베이스를 AWS 클라우드클라우드의 EC2 인스턴스에 있는 Oracle로 마이그레이션합니다.
- 재배포(하이퍼바이저 수준의 리프트 앤 시프트) - 새 하드웨어를 구매하거나, 애플리케이션을 다시 작성하거나, 기존 운영을 수정하지 않고도 인프라를 클라우드로 이동합니다. 온프레미스 플랫폼에서 동일한 플랫폼의 클라우드 서비스로 서버를 마이그레이션합니다. 예: Microsoft Hyper-V 애플리케이션을 로 마이그레이션합니다 AWS.
- 유지(보관) - 소스 환경에 애플리케이션을 유지합니다. 대규모 리팩터링이 필요하고 해당 작업을 나중에 연기하려는 애플리케이션과 비즈니스 차원에서 마이그레이션할 이유가 없어 유지하려는 레거시 애플리케이션이 여기에 포함될 수 있습니다.
- 사용 중지 - 소스 환경에서 더 이상 필요하지 않은 애플리케이션을 폐기하거나 제거합니다.

A

ABAC

[속성 기반 액세스 제어](#)를 참조하세요.

추상화된 서비스

[관리형 서비스](#)를 참조하세요.

ACID

[원자성, 일관성, 격리성, 내구성](#)을 참조하세요.

능동-능동 마이그레이션

양방향 복제 도구 또는 이중 쓰기 작업을 사용하여 소스 데이터베이스와 대상 데이터베이스가 동기화된 상태로 유지되고, 두 데이터베이스 모두 마이그레이션 중 연결 애플리케이션의 트랜잭션을 처리하는 데이터베이스 마이그레이션 방법입니다. 이 방법은 일회성 전환이 필요한 대신 소규모의 제어된 배치로 마이그레이션을 지원합니다. 더 유연하지만 [액티브 패시브 마이그레이션](#)보다 더 많은 작업이 필요합니다.

능동-수동 마이그레이션

소스 데이터베이스와 대상 데이터베이스가 동기화된 상태로 유지되지만 소스 데이터베이스만 연결 애플리케이션의 트랜잭션을 처리하고 데이터는 대상 데이터베이스로 복제되는 데이터베이스 마이그레이션 방법입니다. 대상 데이터베이스는 마이그레이션 중 어떤 트랜잭션도 허용하지 않습니다.

집계 함수

행 그룹에서 작동하고 그룹에 대한 단일 반환 값을 계산하는 SQL 함수입니다. 집계 함수의 예로 SUM 및 MAX가 있습니다.

AI

[인공 지능](#)을 참조하세요.

AI Ops

[인공 지능 운영](#)을 참조하세요.

익명화

데이터세트에서 개인 정보를 영구적으로 삭제하는 프로세스입니다. 익명화는 개인 정보 보호에 도움이 될 수 있습니다. 익명화된 데이터는 더 이상 개인 데이터로 간주되지 않습니다.

안티 패턴

솔루션이 다른 솔루션보다 비생산적이거나 비효율적이거나 덜 효과적이어서 반복되는 문제에 자주 사용되는 솔루션입니다.

애플리케이션 제어

맬웨어로부터 시스템을 보호하기 위해 승인된 애플리케이션만 사용하도록 허용하는 보안 접근 방식입니다.

애플리케이션 포트폴리오

애플리케이션 구축 및 유지 관리 비용과 애플리케이션의 비즈니스 가치를 비롯하여 조직에서 사용하는 각 애플리케이션에 대한 세부 정보 모음입니다. 이 정보는 [포트폴리오 탐색 및 분석 프로세스](#)의 핵심이며 마이그레이션, 현대화 및 최적화할 애플리케이션을 식별하고 우선순위를 정하는 데 도움이 됩니다.

인공 지능

컴퓨터 기술을 사용하여 학습, 문제 해결, 패턴 인식 등 일반적으로 인간과 관련된 인지 기능을 수행하는 것을 전문으로 하는 컴퓨터 과학 분야입니다. 자세한 내용은 [What is Artificial Intelligence?](#)를 참조하십시오.

인공 지능 운영(AIOps)

기계 학습 기법을 사용하여 운영 문제를 해결하고, 운영 인시던트 및 사용자 개입을 줄이고, 서비스 품질을 높이는 프로세스입니다. AWS 마이그레이션 전략에서 AIOps가 사용되는 방법에 대한 자세한 내용은 [운영 통합 가이드](#)를 참조하십시오.

비대칭 암호화

한 쌍의 키, 즉 암호화를 위한 퍼블릭 키와 복호화를 위한 프라이빗 키를 사용하는 암호화 알고리즘입니다. 퍼블릭 키는 복호화에 사용되지 않으므로 공유할 수 있지만 프라이빗 키에 대한 액세스는 엄격히 제한되어야 합니다.

원자성, 일관성, 격리성, 내구성(ACID)

오류, 정전 또는 기타 문제가 발생한 경우에도 데이터베이스의 데이터 유효성과 운영 신뢰성을 보장하는 소프트웨어 속성 세트입니다.

ABAC(속성 기반 액세스 제어)

부서, 직무, 팀 이름 등의 사용자 속성을 기반으로 세분화된 권한을 생성하는 방식입니다. 자세한 내용은 AWS Identity and Access Management (IAM) 설명서의 [용 ABAC AWS](#)를 참조하세요.

신뢰할 수 있는 데이터 소스

가장 신뢰할 수 있는 정보 소스로 간주되는 기본 버전의 데이터를 저장하는 위치입니다. 익명화, 편집 또는 가명화와 같은 데이터 처리 또는 수정의 목적으로 신뢰할 수 있는 데이터 소스의 데이터를 다른 위치로 복사할 수 있습니다.

가용 영역

다른 가용 영역의 장애로부터 격리 AWS 리전 되고 동일한 리전의 다른 가용 영역에 저렴하고 지연 시간이 짧은 네트워크 연결을 제공하는 내의 고유한 위치입니다.

AWS 클라우드 채택 프레임워크(AWS CAF)

조직이 클라우드로 성공적으로 전환 AWS 하기 위한 효율적이고 효과적인 계획을 개발하는 데 도움이 되는 지침 및 모범 사례 프레임워크입니다. AWS CAF는 지침을 비즈니스, 사람, 거버넌스, 플랫폼, 보안 및 운영이라는 6가지 중점 영역으로 구성합니다. 비즈니스, 사람 및 거버넌스 관점은 비즈니스 기술과 프로세스에 초점을 맞추고, 플랫폼, 보안 및 운영 관점은 전문 기술과 프로세스에 중점을 둡니다. 예를 들어, 사람 관점은 인사(HR), 직원 배치 기능 및 인력 관리를 담당하는 이해관계자를 대상으로 합니다. 이러한 관점에서 AWS CAF는 성공적인 클라우드 채택을 위해 조직을 준비하는 데 도움이 되는 인력 개발, 교육 및 커뮤니케이션에 대한 지침을 제공합니다. 자세한 내용은 [AWS CAF 웹사이트](#)와 [AWS CAF 백서](#)를 참조하세요.

AWS 워크로드 검증 프레임워크(AWS WQF)

데이터베이스 마이그레이션 워크로드를 평가하고, 마이그레이션 전략을 권장하고, 작업 견적을 제공하는 도구입니다. AWS WQF는 AWS Schema Conversion Tool (AWS SCT)에 포함되어 있습니다. 데이터베이스 스키마 및 코드 객체, 애플리케이션 코드, 종속성 및 성능 특성을 분석하고 평가 보고서를 제공합니다.

B

악성 봇

개인 또는 조직을 방해하거나 해를 입히기 위한 [봇](#)입니다.

BCP

[비즈니스 연속성 계획](#)을 참조하세요.

동작 그래프

리소스 동작과 시간 경과에 따른 상호 작용에 대한 통합된 대화형 뷰입니다. Amazon Detective에서 동작 그래프를 사용하여 실패한 로그인 시도, 의심스러운 API 직접 호출 및 유사한 작업을 검사할 수 있습니다. 자세한 내용은 Detective 설명서의 [Data in a behavior graph](#)를 참조하십시오.

빅 엔디안 시스템

가장 중요한 바이트를 먼저 저장하는 시스템입니다. [엔디안](#)도 참조하세요.

바이너리 분류

바이너리 결과(가능한 두 클래스 중 하나)를 예측하는 프로세스입니다. 예를 들어, ML 모델이 “이 이메일이 스팸인가요, 스팸이 아닌가요?”, ‘이 제품은 책임가요, 자동차인가요?’ 등의 문제를 예측해야 할 수 있습니다.

블룸 필터

요소가 세트의 멤버인지 여부를 테스트하는 데 사용되는 메모리 효율성이 높은 확률론적 데이터 구조입니다.

블루/그린(Blue/Green) 배포

동일하지만 별개의 두 환경을 생성하는 배포 전략입니다. 하나의 환경(파란색)에서 현재 애플리케이션 버전을 실행하고 새 애플리케이션 버전은 다른 환경(녹색)에서 실행합니다. 이 전략을 사용하면 영향을 최소화하면서 신속하게 롤백할 수 있습니다.

bot

인터넷을 통해 자동화된 태스크를 실행하고 인적 활동이나 상호 작용을 시뮬레이션하는 소프트웨어 애플리케이션입니다. 인터넷에서 정보를 인덱싱하는 웹 크롤러와 같이 유용하거나 이로운 봇도 있습니다. 악성 봇이라고 하는 다른 일부 봇은 개인 또는 조직을 방해하거나 해를 입히기 위한 봇입니다.

봇넷

[맬웨어](#)에 감염되고 봇 허더 또는 봇 운영자와 같은 단일 당사자가 제어하는 [봇](#) 네트워크입니다. 봇넷은 봇의 규모와 봇의 영향 범위를 확대하는 가장 잘 알려진 메커니즘입니다.

브랜치

코드 리포지토리의 포함된 영역입니다. 리포지토리에 생성되는 첫 번째 브랜치가 기본 브랜치입니다. 기존 브랜치에서 새 브랜치를 생성한 다음 새 브랜치에서 기능을 개발하거나 버그를 수정할 수 있습니다. 기능을 구축하기 위해 생성하는 브랜치를 일반적으로 기능 브랜치라고 합니다. 기능을 출시할 준비가 되면 기능 브랜치를 기본 브랜치에 다시 병합합니다. 자세한 내용은 [About branches](#)(GitHub 설명서)를 참조하십시오.

긴급 액세스 권한

예외적인 상황에서 승인된 프로세스를 통해 사용자가 일반적으로 액세스할 권한이 없는데 액세스할 수 있는 빠른 방법입니다. 자세한 내용은 AWS Well-Architected 지침의 [Implement break-glass procedures](#) 지표를 참조하세요.

브라운필드 전략

사용자 환경의 기존 인프라 시스템 아키텍처에 브라운필드 전략을 채택할 때는 현재 시스템 및 인프라의 제약 조건을 중심으로 아키텍처를 설계합니다. 기존 인프라를 확장하는 경우 브라운필드 전략과 [그린필드](#) 전략을 혼합할 수 있습니다.

버퍼 캐시

가장 자주 액세스하는 데이터가 저장되는 메모리 영역입니다.

사업 역량

기업이 가치를 창출하기 위해 하는 일(예: 영업, 고객 서비스 또는 마케팅)입니다. 마이크로서비스 아키텍처 및 개발 결정은 비즈니스 역량에 따라 이루어질 수 있습니다. 자세한 내용은 백서의 [AWS에서 컨테이너화된 마이크로서비스 실행의 비즈니스 역량 중심의 구성화](#) 섹션을 참조하십시오.

비즈니스 연속성 계획(BCP)

대규모 마이그레이션과 같은 중단 이벤트가 운영에 미치는 잠재적 영향을 해결하고 비즈니스가 신속하게 운영을 재개할 수 있도록 지원하는 계획입니다.

C

CAF

[AWS Cloud Adoption Framework](#)를 참조하세요.

카나리 배포

최종 사용자에게 제공하는 느린 증분 릴리스 버전입니다. 확신이 들면 새 버전을 배포하고 현재 버전을 완전히 교체합니다.

CCoE

[클라우드 혁신 센터](#)를 참조하세요.

CDC

[데이터 캡처 변경](#)을 참조하세요.

변경 데이터 캡처(CDC)

데이터베이스 테이블과 같은 데이터 소스의 변경 내용을 추적하고 변경 사항에 대한 메타데이터를 기록하는 프로세스입니다. 대상 시스템의 변경 내용을 감사하거나 복제하여 동기화를 유지하는 등의 다양한 용도로 CDC를 사용할 수 있습니다.

카오스 엔지니어링

시스템의 복원력을 테스트하기 위해 의도적으로 장애나 중단 이벤트를 도입합니다. [AWS Fault Injection Service \(AWS FIS\)](#)를 사용하여 AWS 워크로드에 스트레스를 주고 응답을 평가하는 실험을 수행할 수 있습니다.

CI/CD

[지속적 통합 및 지속적 전송](#)을 참조하세요.

분류

예측을 생성하는 데 도움이 되는 분류 프로세스입니다. 분류 문제에 대한 ML 모델은 이산 값을 예측합니다. 이산 값은 항상 서로 다릅니다. 예를 들어, 모델이 이미지에 자동차가 있는지 여부를 평가해야 할 수 있습니다.

클라이언트측 암호화

대상이 데이터를 AWS 서비스 수신하기 전에 로컬에서 데이터를 암호화합니다.

클라우드 혁신 센터(CCoE)

클라우드 모범 사례 개발, 리소스 동원, 마이그레이션 타임라인 설정, 대규모 혁신을 통한 조직 선도 등 조직 전체에서 클라우드 채택 노력을 추진하는 다분야 팀입니다. 자세한 내용은 AWS 클라우드 엔터프라이즈 전략 블로그의 [CCoE 게시물](#)을 참조하세요.

클라우드 컴퓨팅

원격 데이터 스토리지와 IoT 디바이스 관리에 일반적으로 사용되는 클라우드 기술 클라우드 컴퓨팅은 일반적으로 [엣지 컴퓨팅](#) 기술에 연결되어 있습니다.

클라우드 운영 모델

IT 조직에서 하나 이상의 클라우드 환경을 구축, 성숙화 및 최적화하는 데 사용되는 운영 모델입니다. 자세한 내용은 [클라우드 운영 모델 구축](#)을 참조하십시오.

클라우드 채택 단계

조직이 AWS 클라우드로 마이그레이션할 때 일반적으로 거치는 4단계는 다음과 같습니다.

- 프로젝트 - 개념 증명 및 학습 목적으로 몇 가지 클라우드 관련 프로젝트 실행
- 기반 - 클라우드 채택 확장을 위한 기초 투자(예: 랜딩 존 생성, CCoE 정의, 운영 모델 구축)
- 마이그레이션 - 개별 애플리케이션 마이그레이션
- Re-invention - 제품 및 서비스 최적화와 클라우드 혁신

이러한 단계는 Stephen Orban이 블로그 게시물 [The Journey Toward Cloud-First and the Stages of Adoption](#) on the AWS 클라우드 Enterprise Strategy 블로그에서 정의했습니다. AWS 마이그레이션 전략과 어떤 관련이 있는지에 대한 자세한 내용은 [마이그레이션 준비 가이드](#)를 참조하세요.

CMDB

[구성 관리 데이터베이스](#)를 참조하세요.

코드 리포지토리

소스 코드와 설명서, 샘플, 스크립트 등의 기타 자산이 버전 관리 프로세스를 통해 저장되고 업데이트되는 위치입니다. 일반적인 클라우드 리포지토리로 GitHub 또는 Bitbucket Cloud가 포함됩니다. 코드의 각 버전을 브랜치라고 합니다. 마이크로서비스 구조에서 각 리포지토리는 단일 기능 전용입니다. 단일 CI/CD 파이프라인은 여러 리포지토리를 사용할 수 있습니다.

콜드 캐시

비어 있거나, 제대로 채워지지 않았거나, 오래되었거나 관련 없는 데이터를 포함하는 버퍼 캐시입니다. 주 메모리나 디스크에서 데이터베이스 인스턴스를 읽어야 하기 때문에 성능에 영향을 미치며, 이는 버퍼 캐시에서 읽는 것보다 느립니다.

콜드 데이터

거의 액세스되지 않고 일반적으로 과거 데이터인 데이터. 이런 종류의 데이터를 쿼리할 때는 일반적으로 느린 쿼리가 허용됩니다. 이 데이터를 성능이 낮고 비용이 저렴한 스토리지 계층 또는 클래스로 옮기면 비용을 절감할 수 있습니다.

컴퓨터 비전(CV)

기계 학습을 사용하여 디지털 이미지 및 비디오와 같은 시각적 형식에서 정보를 분석하고 추출하는 [AI](#) 필드입니다. 예를 들어 Amazon SageMaker AI는 CV에 대한 이미지 처리 알고리즘을 제공합니다.

구성 드리프트

워크로드의 경우 구성이 예상되는 상태에서 변경됩니다. 이로 인해 워크로드가 규정을 준수하지 않을 수 있으며, 이는 일반적으로 점진적이고 의도되지 않은 작업입니다.

구성 관리 데이터베이스(CMDB)

하드웨어 및 소프트웨어 구성 요소와 해당 구성을 포함하여 데이터베이스와 해당 IT 환경에 대한 정보를 저장하고 관리하는 리포지토리입니다. 일반적으로 마이그레이션의 포트폴리오 탐색 및 분석 단계에서 CMDB의 데이터를 사용합니다.

규정 준수 팩

규정 준수 및 보안 검사를 사용자 지정하기 위해 조합할 수 있는 AWS Config 규칙 및 수정 작업 모음입니다. YAML 템플릿을 사용하여 적합성 팩을 AWS 계정 및 리전 또는 조직 전체에 단일 엔터티로 배포할 수 있습니다. 자세한 내용은 AWS Config 설명서의 [적합성 팩](#)을 참조하세요.

지속적 통합 및 지속적 전달(CI/CD)

소프트웨어 릴리스 프로세스의 소스, 빌드, 테스트, 스테이징 및 프로덕션 단계를 자동화하는 프로세스입니다. CI/CD는 일반적으로 파이프라인으로 설명됩니다. CI/CD를 통해 프로세스를 자동화하고, 생산성을 높이고, 코드 품질을 개선하고, 더 빠르게 제공할 수 있습니다. 자세한 내용은 [지속적 전달의 이점](#)을 참조하십시오. CD는 지속적 배포를 의미하기도 합니다. 자세한 내용은 [지속적 전달\(Continuous Delivery\)](#)과 [지속적인 개발](#)을 참조하십시오.

CV

[컴퓨터 비전](#)을 참조하세요.

D

저장 데이터

스토리지에 있는 데이터와 같이 네트워크에 고정되어 있는 데이터입니다.

데이터 분류

중요도와 민감도를 기준으로 네트워크의 데이터를 식별하고 분류하는 프로세스입니다. 이 프로세스는 데이터에 대한 적절한 보호 및 보존 제어를 결정하는 데 도움이 되므로 사이버 보안 위험 관리 전략의 중요한 구성 요소입니다. 데이터 분류는 AWS Well-Architected Framework의 보안 원칙 구성 요소입니다. 자세한 내용은 [데이터 분류](#)를 참조하십시오.

데이터 드리프트

프로덕션 데이터와 ML 모델 학습에 사용된 데이터 간의 상당한 차이 또는 시간 경과에 따른 입력 데이터의 의미 있는 변화. 데이터 드리프트는 ML 모델 예측의 전반적인 품질, 정확성 및 공정성을 저하시킬 수 있습니다.

전송 중 데이터

네트워크를 통과하고 있는 데이터입니다. 네트워크 리소스 사이를 이동 중인 데이터를 예로 들 수 있습니다.

데이터 메시

중앙 집중식 관리 및 거버넌스를 통해 분산되고 탈중앙화된 데이터 소유권을 제공하는 아키텍처 프레임워크입니다.

데이터 최소화

꼭 필요한 데이터만 수집하고 처리하는 원칙입니다. 에서 데이터를 최소화하면 개인 정보 보호 위험, 비용 및 분석 탄소 발자국을 줄일 AWS 클라우드 수 있습니다.

데이터 경계

신뢰할 수 있는 자격 증명만 예상 네트워크에서 신뢰할 수 있는 리소스에 액세스하도록 하는 데 도움이 되는 AWS 환경의 예방 가드레일 세트입니다. 자세한 내용은 [데이터 경계 구축을 참조하세요 AWS](#).

데이터 사전 처리

원시 데이터를 ML 모델이 쉽게 구문 분석할 수 있는 형식으로 변환하는 것입니다. 데이터를 사전 처리한다는 것은 특정 열이나 행을 제거하고 누락된 값, 일관성이 없는 값 또는 중복 값을 처리함을 의미할 수 있습니다.

데이터 출처

라이프사이클 전반에 걸쳐 데이터의 출처와 기록을 추적하는 프로세스(예: 데이터 생성, 전송, 저장 방법).

데이터 주체

데이터를 수집 및 처리하는 개인입니다.

데이터 웨어하우스

분석과 같은 비즈니스 인텔리전스를 지원하는 데이터 관리 시스템입니다. 데이터 웨어하우스에는 보통 많은 양의 기록 데이터가 포함되며 일반적으로 쿼리 및 분석에 사용됩니다.

데이터 정의 언어(DDL)

데이터베이스에서 테이블 및 객체의 구조를 만들거나 수정하기 위한 명령문 또는 명령입니다.

데이터베이스 조작 언어(DML)

데이터베이스에서 정보를 수정(삽입, 업데이트 및 삭제)하기 위한 명령문 또는 명령입니다.

DDL

[데이터 정의 언어](#)를 참조하세요.

딥 앙상블

예측을 위해 여러 딥 러닝 모델을 결합하는 것입니다. 딥 앙상블을 사용하여 더 정확한 예측을 얻거나 예측의 불확실성을 추정할 수 있습니다.

딥 러닝

여러 계층의 인공 신경망을 사용하여 입력 데이터와 관심 대상 변수 간의 매핑을 식별하는 ML 하위 분야입니다.

심층 방어

네트워크와 그 안의 데이터 기밀성, 무결성 및 가용성을 보호하기 위해 컴퓨터 네트워크 전체에 일련의 보안 메커니즘과 제어를 신중하게 계층화하는 정보 보안 접근 방식입니다. 이 전략을 채택하면 AWS Organizations 구조의 여러 계층에 여러 제어를 AWS 추가하여 리소스를 보호할 수 있습니다. 예를 들어, 심층 방어 접근 방식은 다단계 인증, 네트워크 세분화 및 암호화를 결합할 수 있습니다.

위임된 관리자

에서 AWS Organizations 호환되는 서비스는 AWS 멤버 계정을 등록하여 조직의 계정을 관리하고 해당 서비스에 대한 권한을 관리할 수 있습니다. 이러한 계정을 해당 서비스의 위임된 관리자라고 합니다. 자세한 내용과 호환되는 서비스 목록은 AWS Organizations 설명서의 [AWS Organizations 와 함께 사용할 수 있는 AWS 서비스](#)를 참조하십시오.

배포

대상 환경에서 애플리케이션, 새 기능 또는 코드 수정 사항을 사용할 수 있도록 하는 프로세스입니다. 배포에는 코드 베이스의 변경 사항을 구현한 다음 애플리케이션 환경에서 해당 코드베이스를 구축하고 실행하는 작업이 포함됩니다.

개발 환경

[환경](#)을 참조하세요.

탐지 제어

이벤트 발생 후 탐지, 기록 및 알림을 수행하도록 설계된 보안 제어입니다. 이러한 제어는 기존의 예방적 제어를 우회한 보안 이벤트를 알리는 2차 방어선입니다. 자세한 내용은 AWS에서 보안 제어 구현의 [탐지 제어](#)를 참조하세요.

개발 가치 흐름 매핑 (DVSM)

소프트웨어 개발 라이프사이클에서 속도와 품질에 부정적인 영향을 미치는 제약 조건을 식별하고 우선 순위를 지정하는 데 사용되는 프로세스입니다. DVSM은 원래 린 제조 방식을 위해 설계된 가치 흐름 매핑 프로세스를 확장합니다. 소프트웨어 개발 프로세스를 통해 가치를 창출하고 이동하는 데 필요한 단계와 팀에 중점을 둡니다.

디지털 트윈

건물, 공장, 산업 장비 또는 생산 라인과 같은 실제 시스템을 가상으로 표현한 것입니다. 디지털 트윈은 예측 유지 보수, 원격 모니터링, 생산 최적화를 지원합니다.

차원 테이블

[스타 스키마](#)에서 팩트 테이블의 정량적 데이터에 대한 데이터 속성을 포함하는 더 작은 테이블을 말합니다. 차원 테이블 속성은 일반적으로 텍스트 필드나 텍스트처럼 동작하는 개별 숫자입니다. 이러한 속성은 보통 쿼리 제약, 필터링 및 결과 세트 레이블 지정에 사용됩니다.

재해

워크로드 또는 시스템이 기본 배포 위치에서 비즈니스 목표를 달성하지 못하게 방해하는 이벤트입니다. 이러한 이벤트는 자연재해, 기술적 오류, 의도하지 않은 구성 오류 또는 멀웨어 공격과 같은 사람의 행동으로 인한 결과일 수 있습니다.

재해 복구(DR)

[재해](#)로 인한 가동 중지 시간 및 데이터 손실을 최소화하기 위해 사용하는 전략 및 프로세스입니다. 자세한 내용은 AWS Well-Architected Framework의 [Disaster Recovery of Workloads on AWS: Recovery in the Cloud](#)를 참조하세요.

DML

[데이터베이스 조작 언어](#)를 참조하세요.

도메인 기반 설계

구성 요소를 각 구성 요소가 제공하는 진화하는 도메인 또는 핵심 비즈니스 목표에 연결하여 복잡한 소프트웨어 시스템을 개발하는 접근 방식입니다. 이 개념은 에릭 에반스에 의해 그의 저서인 도메인 기반 디자인: 소프트웨어 중심의 복잡성 해결(Boston: Addison-Wesley Professional, 2003)에서 소개되었습니다. Strangler Fig 패턴과 함께 도메인 기반 설계를 사용하는 방법에 대한 자세한 내용은 [컨테이너 및 Amazon API Gateway를 사용하여 기존의 Microsoft ASP.NET\(ASMX\) 웹 서비스를 점진적으로 현대화하는 방법](#)을 참조하십시오.

DR

[재해 복구](#)를 참조하세요.

드리프트 감지

기준이 되는 구성과의 편차 추적을 말합니다. 예를 들어 AWS CloudFormation 를 사용하여 [시스템 리소스의 드리프트를 감지](#)하거나 사용하여 AWS Control Tower 거버넌스 요구 사항 준수에 영향을 미칠 수 있는 [랜딩 존의 변경 사항을 감지](#)할 수 있습니다.

DVSM

[개발 가치 흐름 매핑](#)을 참조하세요.

E

EDA

[탐색 데이터 분석](#)을 참조하세요.

EDI

[전자 데이터 교환](#)을 참조하세요.

엣지 컴퓨팅

IoT 네트워크의 엣지에서 스마트 디바이스의 컴퓨팅 성능을 개선하는 기술 엣지 컴퓨팅은 [클라우드 컴퓨팅](#)에 비해 보다 통신 지연 시간을 줄이고 응답 시간을 개선할 수 있습니다.

전자 데이터 교환(EDI)

조직 간 비즈니스 문서의 자동화된 교환을 나타냅니다. 자세한 내용은 [전자 데이터 교환\(EDI\)이란 무엇인가요?](#)를 참조하세요.

암호화

사람이 읽을 수 있는 일반 텍스트 데이터를 사이버텍스트로 변환하는 컴퓨팅 프로세스입니다.

암호화 키

암호화 알고리즘에 의해 생성되는 무작위 비트의 암호화 문자열입니다. 키의 길이는 다양할 수 있으며 각 키는 예측할 수 없고 고유하게 설계되었습니다.

엔디안

컴퓨터 메모리에 바이트가 저장되는 순서입니다. 빅 엔디안 시스템은 가장 중요한 바이트를 먼저 저장합니다. 리틀 엔디안 시스템은 가장 덜 중요한 바이트를 먼저 저장합니다.

엔드포인트

[서비스 엔드포인트](#)를 참조하세요.

엔드포인트 서비스

Virtual Private Cloud(VPC)에서 호스팅하여 다른 사용자와 공유할 수 있는 서비스입니다. 를 사용하여 엔드포인트 서비스를 생성하고 다른 AWS 계정 또는 AWS Identity and Access Management (IAM) 보안 주체에 권한을 AWS PrivateLink 부여할 수 있습니다. 이러한 계정 또는 보안 주체는 인터페이스 VPC 엔드포인트를 생성하여 엔드포인트 서비스에 비공개로 연결할 수 있습니다. 자세한 내용은 Amazon Virtual Private Cloud(VPC) 설명서의 [엔드포인트 서비스 생성](#)을 참조하십시오.

엔터프라이즈 리소스 계획(ERP)

엔터프라이즈의 주요 비즈니스 프로세스(예: 회계, [MES](#), 프로젝트 관리)를 자동화하고 관리하는 시스템입니다.

봉투 암호화

암호화 키를 다른 암호화 키로 암호화하는 프로세스입니다. 자세한 내용은 AWS Key Management Service (AWS KMS) 설명서의 [봉투 암호화](#)를 참조하세요.

환경

실행 중인 애플리케이션의 인스턴스입니다. 다음은 클라우드 컴퓨팅의 일반적인 환경 유형입니다.

- 개발 환경 - 애플리케이션 유지 관리를 담당하는 핵심 팀만 사용할 수 있는 실행 중인 애플리케이션의 인스턴스입니다. 개발 환경은 변경 사항을 상위 환경으로 승격하기 전에 테스트하는 데 사용됩니다. 이러한 유형의 환경을 테스트 환경이라고도 합니다.
- 하위 환경 - 초기 빌드 및 테스트에 사용되는 환경을 비롯한 애플리케이션의 모든 개발 환경입니다.
- 프로덕션 환경 - 최종 사용자가 액세스할 수 있는 실행 중인 애플리케이션의 인스턴스입니다. CI/CD 파이프라인에서 프로덕션 환경이 마지막 배포 환경입니다.
- 상위 환경 - 핵심 개발 팀 이외의 사용자가 액세스할 수 있는 모든 환경입니다. 프로덕션 환경, 프로덕션 이전 환경 및 사용자 수용 테스트를 위한 환경이 여기에 포함될 수 있습니다.

에픽

애자일 방법론에서 작업을 구성하고 우선순위를 정하는 데 도움이 되는 기능적 범주입니다. 에픽은 요구 사항 및 구현 작업에 대한 개괄적인 설명을 제공합니다. 예를 들어, AWS CAF 보안 에픽에는 ID 및 액세스 관리, 탐지 제어, 인프라 보안, 데이터 보호 및 인시던트 대응이 포함됩니다. AWS 마 이그레이션 전략의 에픽에 대한 자세한 내용은 [프로그램 구현 가이드](#)를 참조하십시오.

ERP

[엔터프라이즈 리소스 계획](#)을 참조하세요.

탐색 데이터 분석(EDA)

데이터 세트를 분석하여 주요 특성을 파악하는 프로세스입니다. 데이터를 수집 또는 집계한 다음 초기 조사를 수행하여 패턴을 찾고, 이상을 탐지하고, 가정을 확인합니다. EDA는 요약 통계를 계산하고 데이터 시각화를 생성하여 수행됩니다.

F

팩트 테이블

[스타 스키마](#)의 중앙 테이블입니다. 비즈니스 운영에 대한 정량적 데이터를 저장합니다. 일반적으로 팩트 테이블은 측정값이 있는 열 및 차원 테이블에 대한 외래 키가 있는 열과 같이 두 가지 열 유형을 포함합니다.

빠른 실패

개발 수명 주기를 줄이기 위해 빈번한 증분 테스트를 사용하는 철학입니다. 애자일 접근 방식의 핵심입니다.

장애 격리 경계

에서 장애의 영향을 제한하고 워크로드의 복원력을 개선하는 데 도움이 되는 가용 영역, AWS 리전 컨트롤 플레인 또는 데이터 플레인과 같은 AWS 클라우드경계입니다. 자세한 내용은 [AWS 장애 격리 경계](#)를 참조하세요.

기능 브랜치

[브랜치](#)를 참조하세요.

기능

예측에 사용하는 입력 데이터입니다. 예를 들어, 제조 환경에서 기능은 제조 라인에서 주기적으로 캡처되는 이미지일 수 있습니다.

기능 중요도

모델의 예측에 특성이 얼마나 중요한지를 나타냅니다. 이는 일반적으로 SHAP(Shapley Additive Descriptions) 및 통합 그래디언트와 같은 다양한 기법을 통해 계산할 수 있는 수치 점수로 표현됩니다. 자세한 내용은 [기계 학습 모델 해석 가능성을 참조하세요 AWS](#).

기능 변환

추가 소스로 데이터를 보강하거나, 값을 조정하거나, 단일 데이터 필드에서 여러 정보 세트를 추출하는 등 ML 프로세스를 위해 데이터를 최적화하는 것입니다. 이를 통해 ML 모델이 데이터를 활용

할 수 있습니다. 예를 들어, 날짜 '2021-05-27 00:15:37'을 '2021년', '5월', '목', '15일'로 분류하면 학습 알고리즘이 다양한 데이터 구성 요소와 관련된 미묘한 패턴을 학습하는 데 도움이 됩니다.

퓨샷 프롬프팅

유사한 태스크를 수행하도록 요청하기 전에 [LLM](#)에 태스크와 원하는 출력을 보여주는 몇 가지 예제를 제공합니다. 이 기법은 모델이 프롬프트에 포함된 예제(샷)에서 학습하는 컨텍스트 내 학습을 적용합니다. 퓨샷 프롬프팅은 특정 형식 지정, 추론 또는 분야별 지식이 필요한 태스크에 효과적일 수 있습니다. [제로샷 프롬프팅](#)도 참조하세요.

FGAC

[세분화된 액세스 제어](#)를 참조하세요.

세분화된 액세스 제어(FGAC)

여러 조건을 사용하여 액세스 요청을 허용하거나 거부합니다.

플래시컷 마이그레이션

단계적 접근 방식을 사용하는 대신 [변경 데이터 캡처](#)를 통해 지속적 데이터 복제를 사용하여 최단 시간에 데이터를 마이그레이션하는 데이터베이스 마이그레이션 방법입니다. 목표는 가동 중지 시간을 최소화하는 것입니다.

FM

[파운데이션 모델](#)을 참조하세요.

파운데이션 모델(FM)

일반화되고 레이블이 지정되지 않은 데이터의 대규모 데이터세트에서 훈련된 대규모 딥 러닝 신경망입니다. FM은 언어 이해, 텍스트 및 이미지 생성, 자연어 대화와 같은 다양한 일반 태스크를 수행할 수 있습니다. 자세한 내용은 [파운데이션 모델이란?](#)을 참조하세요.

G

생성형 AI

대량의 데이터에서 훈련되었으며 간단한 텍스트 프롬프트를 사용하여 이미지, 비디오, 텍스트, 오디오와 같은 새 콘텐츠와 아티팩트를 생성할 수 있는 [AI](#) 모델의 하위 세트입니다. 자세한 내용은 [생성형 AI란 무엇인가요?](#)를 참조하세요.

지리적 차단

[지리적 제한](#)을 참조하세요.

지리적 제한(지리적 차단)

Amazon CloudFront에서 특정 국가의 사용자가 콘텐츠 배포에 액세스하지 못하도록 하는 옵션입니다. 허용 목록 또는 차단 목록을 사용하여 승인된 국가와 차단된 국가를 지정할 수 있습니다. 자세한 내용은 CloudFront 설명서의 [콘텐츠의 지리적 배포 제한](#)을 참조하십시오.

Gitflow 워크플로

하위 환경과 상위 환경이 소스 코드 리포지토리의 서로 다른 브랜치를 사용하는 방식입니다. Gitflow 워크플로는 레거시로 간주되며 [트렁크 기반 워크플로](#)는 선호되는 현대적 접근 방식입니다.

골든 이미지

시스템 또는 소프트웨어의 새 인스턴스를 배포하기 위한 템플릿으로 사용되는 해당 시스템 또는 소프트웨어의 스냅샷입니다. 예를 들어 제조 분야에서는 골든 이미지를 사용하여 여러 디바이스에서 소프트웨어를 프로비저닝할 수 있으며 이를 통해 디바이스 제조 작업의 속도, 확장성 및 생산성을 개선할 수 있습니다.

브라운필드 전략

새로운 환경에서 기존 인프라의 부재 시스템 아키텍처에 대한 그린필드 전략을 채택할 때 [브라운필드](#)라고도 하는 기존 인프라와의 호환성 제한 없이 모든 새로운 기술을 선택할 수 있습니다. 기존 인프라를 확장하는 경우 브라운필드 전략과 그린필드 전략을 혼합할 수 있습니다.

가드레일

조직 단위(OU) 전체에서 리소스, 정책 및 규정 준수를 관리하는 데 도움이 되는 중요 규칙입니다. 예방 가드레일은 규정 준수 표준에 부합하도록 정책을 시행하며, 서비스 제어 정책과 IAM 권한 경계를 사용하여 구현됩니다. 탐지 가드레일은 정책 위반 및 규정 준수 문제를 감지하고 해결을 위한 알림을 생성하며, 이는 AWS Config, Amazon GuardDuty, AWS Security Hub CSPM, , AWS Trusted Advisor, Amazon Inspector 및 사용자 지정 AWS Lambda 검사를 사용하여 구현됩니다.

H

HA

[고가용성](#)을 참조하세요.

이기종 데이터베이스 마이그레이션

다른 데이터베이스 엔진을 사용하는 대상 데이터베이스로 소스 데이터베이스 마이그레이션(예: Oracle에서 Amazon Aurora로) 이기종 마이그레이션은 일반적으로 리아키텍트 작업의 일부이며 스

키마를 변환하는 것은 복잡한 작업일 수 있습니다. AWS 는 스키마 변환에 도움이 되는 [AWS SCT](#)를 제공합니다.

높은 가용성(HA)

문제나 재해 발생 시 개입 없이 지속적으로 운영할 수 있는 워크로드의 능력. HA 시스템은 자동으로 장애 조치되고, 지속적으로 고품질 성능을 제공하고, 성능에 미치는 영향을 최소화하면서 다양한 부하와 장애를 처리하도록 설계되었습니다.

히스토리언 현대화

제조 산업의 요구 사항을 더 잘 충족하도록 운영 기술(OT) 시스템을 현대화하고 업그레이드하는 데 사용되는 접근 방식입니다. 히스토리언은 공장의 다양한 출처에서 데이터를 수집하고 저장하는 데 사용되는 일종의 데이터베이스입니다.

홀드아웃 데이터

[기계 학습](#) 모델을 훈련하는 데 사용되는 데이터세트에서 보류되는 레이블이 지정된 기록 데이터의 일부입니다. 홀드아웃 데이터를 사용하여 모델 예측을 홀드아웃 데이터와 비교해 모델 성능을 평가할 수 있습니다.

동종 데이터베이스 마이그레이션

동일한 데이터베이스 엔진을 공유하는 대상 데이터베이스로 소스 데이터베이스 마이그레이션(예: Microsoft SQL Server에서 Amazon RDS for SQL Server로) 동종 마이그레이션은 일반적으로 리호스팅 또는 리플랫폼 작업의 일부입니다. 네이티브 데이터베이스 유틸리티를 사용하여 스키마를 마이그레이션할 수 있습니다.

핫 데이터

자주 액세스하는 데이터(예: 실시간 데이터 또는 최근 번역 데이터). 일반적으로 이 데이터에는 빠른 쿼리 응답을 제공하기 위한 고성능 스토리지 계층 또는 클래스가 필요합니다.

핫픽스

프로덕션 환경의 중요한 문제를 해결하기 위한 긴급 수정입니다. 핫픽스는 긴급하기 때문에 일반적인 DevOps 릴리스 워크플로 외부에서 실행됩니다.

하이퍼케어 기간

전환 직후 마이그레이션 팀이 문제를 해결하기 위해 클라우드에서 마이그레이션된 애플리케이션을 관리하고 모니터링하는 기간입니다. 일반적으로 이 기간은 1~4일입니다. 하이퍼케어 기간이 끝나면 마이그레이션 팀은 일반적으로 애플리케이션에 대한 책임을 클라우드 운영 팀에 넘깁니다.

I

IaC

[코드형 인프라](#)를 참조하세요.

자격 증명 기반 정책

AWS 클라우드 환경 내에서 권한을 정의하는 하나 이상의 IAM 보안 주체에 연결된 정책입니다.

유휴 애플리케이션

90일 동안 평균 CPU 및 메모리 사용량이 5~20%인 애플리케이션입니다. 마이그레이션 프로젝트에 서는 이러한 애플리케이션을 사용 중지하거나 온프레미스에 유지하는 것이 일반적입니다.

IIoT

[산업용 사물 인터넷](#)을 참조하세요.

변경 불가능한 인프라

기존 인프라를 업데이트, 패치 또는 수정하는 대신 프로덕션 워크로드에 대한 새 인프라를 배포하 는 모델입니다. 변경 불가능한 인프라는 [변경 가능한 인프라](#)보다 본질적으로 더 일관되고 안정적이 며 예측 가능합니다. 자세한 내용은 AWS Well-Architected Framework의 [변경 불가능한 인프라를 사용하여 배포](#) 모범 사례를 참조하세요.

인바운드(수신) VPC

AWS 다중 계정 아키텍처에서 애플리케이션 외부에서 네트워크 연결을 수락, 검사 및 라우팅하는 VPC입니다. [AWS Security Reference Architecture](#)에서는 애플리케이션과 더 넓은 인터넷 간의 양 방향 인터페이스를 보호하기 위해 인바운드, 아웃바운드 및 검사 VPC로 네트워크 계정을 설정할 것을 권장합니다.

증분 마이그레이션

한 번에 전체 전환을 수행하는 대신 애플리케이션을 조금씩 마이그레이션하는 전환 전략입니다. 예 를 들어, 처음에는 소수의 마이크로서비스나 사용자만 새 시스템으로 이동할 수 있습니다. 모든 것 이 제대로 작동하는지 확인한 후에는 레거시 시스템을 폐기할 수 있을 때까지 추가 마이크로서비스 또는 사용자를 점진적으로 이동할 수 있습니다. 이 전략을 사용하면 대규모 마이그레이션과 관련된 위험을 줄일 수 있습니다.

Industry 4.0

연결성, 실시간 데이터, 자동화, 분석 및 AI/ML의 발전을 통해 제조 프로세스의 현대화를 나타내기 위해 2016년에 [Klaus Schwab](#)에서 도입한 용어입니다.

인프라

애플리케이션의 환경 내에 포함된 모든 리소스와 자산입니다.

코드형 인프라(IaC)

구성 파일 세트를 통해 애플리케이션의 인프라를 프로비저닝하고 관리하는 프로세스입니다. IaC는 새로운 환경의 반복 가능성, 신뢰성 및 일관성을 위해 인프라 관리를 중앙 집중화하고, 리소스를 표준화하고, 빠르게 확장할 수 있도록 설계되었습니다.

산업용 사물 인터넷(IIoT)

제조, 에너지, 자동차, 의료, 생명과학, 농업 등의 산업 부문에서 인터넷에 연결된 센서 및 디바이스의 사용 자세한 내용은 [산업용 사물 인터넷\(IoT\) 디지털 트랜스포메이션 전략 구축](#)을 참조하십시오.

검사 VPC

AWS 다중 계정 아키텍처에서는 VPC(동일하거나 다른 AWS 리전), 인터넷 및 온프레미스 네트워크 간의 네트워크 트래픽 검사를 관리하는 중앙 집중식 VPCs입니다. [AWS Security Reference Architecture](#)에서는 애플리케이션과 더 넓은 인터넷 간의 양방향 인터페이스를 보호하기 위해 인바운드, 아웃바운드 및 검사 VPC로 네트워크 계정을 설정할 것을 권장합니다.

사물 인터넷(IoT)

인터넷이나 로컬 통신 네트워크를 통해 다른 디바이스 및 시스템과 통신하는 센서 또는 프로세서가 내장된 연결된 물리적 객체의 네트워크 자세한 내용은 [IoT란?](#)을 참조하십시오.

해석력

모델의 예측이 입력에 따라 어떻게 달라지는지를 사람이 이해할 수 있는 정도를 설명하는 기계 학습 모델의 특성입니다. 자세한 내용은 [기계 학습 모델 해석 가능성을 참조하세요 AWS](#).

IoT

[사물 인터넷](#)을 참조하세요.

IT 정보 라이브러리(ITIL)

IT 서비스를 제공하고 이러한 서비스를 비즈니스 요구 사항에 맞게 조정하기 위한 일련의 모범 사례 ITIL은 ITSM의 기반을 제공합니다.

IT 서비스 관리(ITSM)

조직의 IT 서비스 설계, 구현, 관리 및 지원과 관련된 활동 클라우드 운영을 ITSM 도구와 통합하는 방법에 대한 자세한 내용은 [운영 통합 가이드](#)를 참조하십시오.

ITIL

[IT 정보 라이브러리](#)를 참조하세요.

ITSM

[IT 서비스 관리](#)를 참조하세요.

L

레이블 기반 액세스 제어(LBAC)

사용자 및 데이터 자체에 각각 보안 레이블 값을 명시적으로 할당하는 필수 액세스 제어(MAC)를 구현한 것입니다. 사용자 보안 레이블과 데이터 보안 레이블 간의 교차 부분에 따라 사용자가 볼 수 있는 행과 열이 결정됩니다.

랜딩 존

랜딩 존은 확장 가능하고 안전한 잘 설계된 다중 계정 AWS 환경입니다. 조직은 여기에서부터 보안 및 인프라 환경에 대한 확신을 가지고 워크로드와 애플리케이션을 신속하게 시작하고 배포할 수 있습니다. 랜딩 존에 대한 자세한 내용은 [안전하고 확장 가능한 다중 계정 AWS 환경 설정](#)을 참조하십시오.

대규모 언어 모델(LLM)

방대한 양의 데이터에서 사전 훈련된 딥 러닝 [AI](#) 모델입니다. LLM은 질문에 대한 답변, 문서 요약, 텍스트를 다른 언어로 번역, 문장 완성과 같은 여러 태스크를 수행할 수 있습니다. 자세한 내용은 [대규모 언어 모델\(LLM\)이란 무엇인가요?](#)를 참조하세요.

대규모 마이그레이션

300대 이상의 서버 마이그레이션입니다.

LBAC

[레이블 기반 액세스 제어](#)를 참조하세요.

최소 권한

작업을 수행하는 데 필요한 최소 권한을 부여하는 보안 모범 사례입니다. 자세한 내용은 IAM 설명서의 [최소 권한 적용](#)을 참조하십시오.

리프트 앤드 시프트

[7R](#)을 참조하세요.

리틀 엔디안 시스템

가장 덜 중요한 바이트를 먼저 저장하는 시스템입니다. [엔디안](#)도 참조하세요.

LLM

[대규모 언어 모델](#)을 참조하세요.

하위 환경

[환경](#)을 참조하세요.

M

기계 학습(ML)

패턴 인식 및 학습에 알고리즘과 기법을 사용하는 인공지능의 한 유형입니다. ML은 사물 인터넷 (IoT) 데이터와 같은 기록된 데이터를 분석하고 학습하여 패턴을 기반으로 통계 모델을 생성합니다. 자세한 내용은 [기계 학습](#)을 참조하십시오.

기본 브랜치

[브랜치](#)를 참조하세요.

맬웨어

컴퓨터 보안 또는 프라이버시를 위협하도록 설계된 소프트웨어입니다. 맬웨어는 컴퓨터 시스템을 방해하거나 민감한 정보를 유출하거나 무단 액세스 권한을 확보할 수 있습니다. 맬웨어의 예로 바이러스, 웜, 랜섬웨어, 트로이 목마, 스파이웨어, 키로거 등이 있습니다.

관리형 서비스

AWS 서비스는 인프라 계층, 운영 체제 및 플랫폼을 AWS 운영하고, 사용자는 엔드포인트에 액세스하여 데이터를 저장하고 검색합니다. 관리형 서비스의 예로 Amazon Simple Storage Service(Amazon S3) 및 Amazon DynamoDB가 있습니다. 이를 추상화된 서비스라고도 합니다.

제조 실행 시스템(MES)

원자재를 생산 현장에서 완제품으로 변환하는 생산 프로세스를 추적, 모니터링, 문서화 및 제어하기 위한 소프트웨어 시스템입니다.

MAP

[Migration Acceleration Program](#)을 참조하세요.

메커니즘

도구를 생성하고 도구 채택을 유도한 다음 조정을 위해 결과를 검사하는 전체 프로세스입니다. 메커니즘은 작동 시 자체적으로 강화하고 개선하는 주기입니다. 자세한 내용은 AWS Well-Architected Framework의 [메커니즘 구축](#)을 참조하세요.

멤버 계정

조직의 일부인 관리 계정을 AWS 계정 제외한 모든 계정. AWS Organizations 하나의 계정은 한 번에 하나의 조직 멤버만 될 수 있습니다.

MES

[제조 실행 시스템](#)을 참조하세요.

메시지 큐 원격 분석 전송(MQTT)

리소스 제약이 있는 [IoT](#) 디바이스에 대한 [게시 및 구독](#) 패턴을 기반으로 하는 경량 Machine-to-Machine(M2M) 통신 프로토콜입니다.

마이크로서비스

잘 정의된 API를 통해 통신하고 일반적으로 소규모 자체 팀이 소유하는 소규모 독립 서비스입니다. 예를 들어, 보험 시스템에는 영업, 마케팅 등의 비즈니스 역량이나 구매, 청구, 분석 등의 하위 영역에 매핑되는 마이크로 서비스가 포함될 수 있습니다. 마이크로서비스의 이점으로 민첩성, 유연한 확장, 손쉬운 배포, 재사용 가능한 코드, 복원력 등이 있습니다. 자세한 내용은 [AWS 서버리스 서비스를 사용하여 마이크로서비스 통합을 참조하세요](#).

마이크로서비스 아키텍처

각 애플리케이션 프로세스를 마이크로서비스로 실행하는 독립 구성 요소를 사용하여 애플리케이션을 구축하는 접근 방식입니다. 이러한 마이크로서비스는 경량 API를 사용하여 잘 정의된 인터페이스를 통해 통신합니다. 애플리케이션의 특정 기능에 대한 수요에 맞게 이 아키텍처의 각 마이크로 서비스를 업데이트, 배포 및 조정할 수 있습니다. 자세한 내용은 [에서 마이크로서비스 구현을 참조하세요 AWS](#).

Migration Acceleration Program(MAP)

조직이 클라우드로 전환하기 위한 강력한 운영 기반을 구축하고 초기 마이그레이션 비용을 상쇄하는 데 도움이 되는 컨설팅 지원, 교육 및 서비스를 제공하는 AWS 프로그램입니다. MAP에는 레거시 마이그레이션을 체계적인 방식으로 실행하기 위한 마이그레이션 방법론과 일반적인 마이그레이션 시나리오를 자동화하고 가속화하는 도구 세트가 포함되어 있습니다.

대규모 마이그레이션

애플리케이션 포트폴리오의 대다수를 웨이브를 통해 클라우드로 이동하는 프로세스로, 각 웨이브에서 더 많은 애플리케이션이 더 빠른 속도로 이동합니다. 이 단계에서는 이전 단계에서 배운 모범 사례와 교훈을 사용하여 팀, 도구 및 프로세스의 마이그레이션 팩토리를 구현하여 자동화 및 민첩한 제공을 통해 워크로드 마이그레이션을 간소화합니다. 이것은 [AWS 마이그레이션 전략](#)의 세 번째 단계입니다.

마이그레이션 팩토리

자동화되고 민첩한 접근 방식을 통해 워크로드 마이그레이션을 간소화하는 다기능 팀입니다. 마이그레이션 팩토리 팀에는 일반적으로 스프린트에서 일하는 운영, 비즈니스 분석가 및 소유자, 마이그레이션 엔지니어, 개발자, DevOps 전문가가 포함됩니다. 엔터프라이즈 애플리케이션 포트폴리오의 20~50%는 공장 접근 방식으로 최적화할 수 있는 반복되는 패턴으로 구성되어 있습니다. 자세한 내용은 이 콘텐츠 세트의 [클라우드 마이그레이션 팩토리 가이드](#)와 [마이그레이션 팩토리에 대한 설명](#)을 참조하십시오.

마이그레이션 메타데이터

마이그레이션을 완료하는 데 필요한 애플리케이션 및 서버에 대한 정보 각 마이그레이션 패턴에는 서로 다른 마이그레이션 메타데이터 세트가 필요합니다. 마이그레이션 메타데이터의 예로는 대상 서브넷, 보안 그룹 및 AWS 계정이 있습니다.

마이그레이션 패턴

사용되는 마이그레이션 전략, 마이그레이션 대상, 마이그레이션 애플리케이션 또는 서비스를 자세히 설명하는 반복 가능한 마이그레이션 작업입니다. 예: AWS Application Migration Service를 사용하여 Amazon EC2로 마이그레이션을 리호스팅합니다.

Migration Portfolio Assessment(MPA)

AWS 클라우드로 마이그레이션하는 비즈니스 사례를 검증하기 위한 정보를 제공하는 온라인 도구입니다. MPA는 상세한 포트폴리오 평가(서버 적정 규모 조정, 가격 책정, TCO 비교, 마이그레이션 비용 분석)와 마이그레이션 계획(애플리케이션 데이터 분석 및 데이터 수집, 애플리케이션 그룹화, 마이그레이션 우선순위 지정, 웨이브 계획)을 제공합니다. [MPA 도구](#)(로그인 필요)는 모든 AWS 컨설턴트와 APN 파트너 컨설턴트가 무료로 사용할 수 있습니다.

마이그레이션 준비 상태 평가(MRA)

AWS CAF를 사용하여 조직의 클라우드 준비 상태에 대한 인사이트를 얻고, 강점과 약점을 식별하고, 식별된 격차를 해소하기 위한 행동 계획을 수립하는 프로세스입니다. 자세한 내용은 [마이그레이션 준비 가이드](#)를 참조하십시오. MRA는 [AWS 마이그레이션 전략](#)의 첫 번째 단계입니다.

마이그레이션 전략

워크로드를 AWS 클라우드로 마이그레이션하는 데 사용되는 접근 방식입니다. 자세한 내용은 이 용어집의 [7R 항목](#)과 [조직을 동원하여 대규모 마이그레이션 가속화](#)를 참조하세요.

ML

[기계 학습](#)을 참조하세요.

현대화

비용을 절감하고 효율성을 높이고 혁신을 활용하기 위해 구식(레거시 또는 모놀리식) 애플리케이션과 해당 인프라를 클라우드의 민첩하고 탄력적이고 가용성이 높은 시스템으로 전환하는 것입니다. 자세한 내용은 [AWS 클라우드에서 애플리케이션을 현대화하기 위한 전략](#)을 참조하세요.

현대화 준비 상태 평가

조직 애플리케이션의 현대화 준비 상태를 파악하고, 이점, 위험 및 종속성을 식별하고, 조직이 해당 애플리케이션의 향후 상태를 얼마나 잘 지원할 수 있는지를 확인하는 데 도움이 되는 평가입니다. 평가 결과는 대상 아키텍처의 청사진, 현대화 프로세스의 개발 단계와 마일스톤을 자세히 설명하는 로드맵 및 파악된 격차를 해소하기 위한 실행 계획입니다. 자세한 내용은 [AWS 클라우드에서 애플리케이션의 현대화 준비 상태 평가](#)를 참조하세요.

모놀리식 애플리케이션(모놀리식 유형)

긴밀하게 연결된 프로세스를 사용하여 단일 서비스로 실행되는 애플리케이션입니다. 모놀리식 애플리케이션에는 몇 가지 단점이 있습니다. 한 애플리케이션 기능에 대한 수요가 급증하면 전체 아키텍처 규모를 조정해야 합니다. 코드 베이스가 커지면 모놀리식 애플리케이션의 기능을 추가하거나 개선하는 것도 더 복잡해집니다. 이러한 문제를 해결하기 위해 마이크로서비스 아키텍처를 사용할 수 있습니다. 자세한 내용은 [마이크로서비스로 모놀리식 유형 분해](#)를 참조하십시오.

MPA

[Migration Portfolio Assessment](#)를 참조하세요.

MQTT

[메시지 큐 원격 분석 전송](#)을 참조하세요.

멀티클래스 분류

여러 클래스에 대한 예측(2개 이상의 결과 중 하나 예측)을 생성하는 데 도움이 되는 프로세스입니다. 예를 들어, ML 모델이 '이 제품은 책인가요, 자동차인가요, 휴대폰인가요?' 또는 '이 고객이 가장 관심을 갖는 제품 범주는 무엇인가요?'라고 물을 수 있습니다.

변경 가능한 인프라

프로덕션 워크로드에 대한 기존 인프라를 업데이트하고 수정하는 모델입니다. 일관성, 신뢰성 및 예측 가능성을 높이기 위해 AWS Well-Architected Framework에서는 [변경 불가능한 인프라](#)를 모범 사례로 사용할 것을 권장합니다.

O

OAC

[오리진 액세스 제어](#)를 참조하세요.

OAI

[오리진 액세스 ID](#)를 참조하세요.

OCM

[조직 변경 관리](#)를 참조하세요.

오프라인 마이그레이션

마이그레이션 프로세스 중 소스 워크로드가 중단되는 마이그레이션 방법입니다. 이 방법은 가동 중지 증가를 수반하며 일반적으로 작고 중요하지 않은 워크로드에 사용됩니다.

O

[운영 통합](#)을 참조하세요.

OLA

[운영 수준 계약](#)을 참조하세요.

온라인 마이그레이션

소스 워크로드를 오프라인 상태로 전환하지 않고 대상 시스템에 복사하는 마이그레이션 방법입니다. 워크로드에 연결된 애플리케이션은 마이그레이션 중에도 계속 작동할 수 있습니다. 이 방법은 가동 중지 차단 또는 최소화를 수반하며 일반적으로 중요한 프로덕션 워크로드에 사용됩니다.

OPC-UA

[Open Process Communications - Unified Architecture\(OPC-UA\)](#)를 참조하세요.

Open Process Communications - Unified Architecture(OPC-UA)

산업 자동화를 위한 Machine-to-Machine(M2M) 통신 프로토콜입니다. OPC-UA는 데이터 암호화, 인증 및 권한 부여 체계에 관한 상호 운용성 표준을 제공합니다.

운영 수준 협약(OLA)

서비스 수준에 관한 계약(SLA)을 지원하기 위해 직무 IT 그룹이 서로에게 제공하기로 약속한 내용을 명확히 하는 계약입니다.

운영 준비 상태 검토(ORR)

인시던트 및 잠재적 장애의 범위를 이해, 평가 또는 예방하거나 줄이는 데 도움이 되는 질문 체크리스트 및 관련 모범 사례입니다. 자세한 내용은 AWS Well-Architected Framework의 [운영 준비 상태 검토\(ORR\)](#)를 참조하세요.

운영 기술(OT)

물리적 환경에서 작동하여 산업 운영, 장비 및 인프라를 제어하는 하드웨어 및 소프트웨어 시스템입니다. 제조 분야에서 OT 및 정보 기술(IT) 시스템의 통합은 [Industry 4.0](#) 트랜스포메이션의 주요 중점 사항입니다.

운영 통합(OI)

클라우드에서 운영을 현대화하는 프로세스로 준비 계획, 자동화 및 통합을 수반합니다. 자세한 내용은 [운영 통합 가이드](#)를 참조하십시오.

조직 트레일

조직 AWS 계정 내 모든에 대한 모든 이벤트를 로깅 AWS CloudTrail 하는에서 생성된 추적입니다 AWS Organizations. 이 트레일은 조직에 속한 각 AWS 계정에 생성되고 각 계정의 활동을 추적합니다. 자세한 내용은 CloudTrail 설명서의 [Creating a trail for an organization](#)을 참조하십시오.

조직 변경 관리(OCM)

사람, 문화 및 리더십 관점에서 중대하고 파괴적인 비즈니스 혁신을 관리하기 위한 프레임워크입니다. OCM은 변화 채택을 가속화하고, 과도기적 문제를 해결하고, 문화 및 조직적 변화를 주도함으로써 조직이 새로운 시스템 및 전략을 준비하고 전환할 수 있도록 지원합니다. AWS 마이그레이션 전략에서는 클라우드 채택 프로젝트에 필요한 변경 속도 때문에이 프레임워크를 인력 가속화라고 합니다. 자세한 내용은 [사용 가이드](#)를 참조하십시오.

오리진 액세스 제어(OAC)

CloudFront에서 Amazon Simple Storage Service(S3) 콘텐츠를 보호하기 위해 액세스를 제한하는 고급 옵션입니다. OAC는 AWS KMS (SSE-KMS)를 사용한 모든 서버 측 암호화 AWS 리전와 S3 버킷에 대한 동적 PUT 및 DELETE 요청에서 모든 S3 버킷을 지원합니다.

오리진 액세스 ID(OAI)

CloudFront에서 Amazon S3 콘텐츠를 보호하기 위해 액세스를 제한하는 옵션입니다. OAI를 사용하면 CloudFront는 Amazon S3가 인증할 수 있는 보안 주체를 생성합니다. 인증된 보안 주체는 특

정 CloudFront 배포를 통해서만 S3 버킷의 콘텐츠에 액세스할 수 있습니다. 더 세분화되고 향상된 액세스 제어를 제공하는 [OAC](#)도 참조하십시오.

ORR

[운영 준비 상태 검토](#)를 참조하세요.

OT

[운영 기술](#)을 참조하세요.

아웃바운드(송신) VPC

AWS 다중 계정 아키텍처에서 애플리케이션 내에서 시작된 네트워크 연결을 처리하는 VPC입니다. [AWS Security Reference Architecture](#)에서는 애플리케이션과 더 넓은 인터넷 간의 양방향 인터페이스를 보호하기 위해 인바운드, 아웃바운드 및 검사 VPC로 네트워크 계정을 설정할 것을 권장합니다.

P

권한 경계

사용자나 역할이 가질 수 있는 최대 권한을 설정하기 위해 IAM 보안 주체에 연결되는 IAM 관리 정책입니다. 자세한 내용은 IAM 설명서의 [권한 경계](#)를 참조하십시오.

개인 식별 정보(PII)

직접 보거나 다른 관련 데이터와 함께 짝을 지을 때 개인의 신원을 합리적으로 추론하는 데 사용할 수 있는 정보입니다. PII의 예로는 이름, 주소, 연락처 정보 등이 있습니다.

PII

[개인 식별 정보](#)를 참조하세요.

플레이북

클라우드에서 핵심 운영 기능을 제공하는 등 마이그레이션과 관련된 작업을 캡처하는 일련의 사전 정의된 단계입니다. 플레이북은 스크립트, 자동화된 런북 또는 현대화된 환경을 운영하는 데 필요한 프로세스나 단계 요약의 형태를 취할 수 있습니다.

PLC

[프로그래밍 가능 로직 컨트롤러](#)를 참조하세요.

PLM

[제품 수명 주기 관리](#)를 참조하세요.

정책

권한 정의([ID 기반 정책](#) 참조), 액세스 조건 지정([리소스 기반 정책](#) 참조), AWS Organizations 내 조직의 모든 계정에 대한 최대 권한 정의([서비스 제어 정책](#) 참조)와 같은 작업을 수행할 수 있는 객체입니다.

다국어 지속성

데이터 액세스 패턴 및 기타 요구 사항을 기반으로 독립적으로 마이크로서비스의 데이터 스토리지 기술 선택. 마이크로서비스가 동일한 데이터 스토리지 기술을 사용하는 경우 구현 문제가 발생하거나 성능이 저하될 수 있습니다. 요구 사항에 가장 적합한 데이터 저장소를 사용하면 마이크로서비스를 더 쉽게 구현하고 성능과 확장성을 높일 수 있습니다.

포트폴리오 평가

마이그레이션을 계획하기 위해 애플리케이션 포트폴리오를 검색 및 분석하고 우선순위를 정하는 프로세스입니다. 자세한 내용은 [마이그레이션 준비 상태 평가](#)를 참조하십시오.

조건자

보통 WHERE 절에 있는 true 또는 false를 반환하는 쿼리 조건입니다.

푸시다운 조건자

전송 전에 쿼리의 데이터를 필터링하는 데이터베이스 쿼리 최적화 기법입니다. 이렇게 하면 관계형 데이터베이스에서 검색하고 처리해야 하는 데이터의 양이 줄고 쿼리 성능이 향상됩니다.

예방적 제어

이벤트 발생을 방지하도록 설계된 보안 제어입니다. 이 제어는 네트워크에 대한 무단 액세스나 원치 않는 변경을 방지하는 데 도움이 되는 1차 방어선입니다. 자세한 내용은 Implementing security controls on AWS의 [Preventative controls](#)를 참조하십시오.

보안 주체

작업을 수행하고 리소스에 액세스할 수 있는 AWS 있는의 엔터티입니다. 이 엔터티는 일반적으로 , AWS 계정 IAM 역할 또는 사용자의 루트 사용자입니다. 자세한 내용은 IAM 설명서의 [역할 용어 및 개념](#)의 보안 주체를 참조하십시오.

개인 정보 보호 중심 설계

전체 개발 프로세스에서 개인 정보를 고려하는 시스템 엔지니어링에서의 접근 방식입니다.

프라이빗 호스팅 영역

Amazon Route 53에서 하나 이상의 VPC 내 도메인과 하위 도메인에 대한 DNS 쿼리에 응답하는 방법에 대한 정보가 담긴 컨테이너입니다. 자세한 내용은 Route 53 설명서의 [프라이빗 호스팅 영역 작업](#)을 참조하십시오.

선제적 제어

규정 미준수 리소스의 배포를 방지하도록 설계된 [보안 제어](#)입니다. 이러한 제어는 리소스를 프로비저닝하기 전에 리소스를 스캔합니다. 리소스가 제어를 준수하지 않으면 프로비저닝되지 않습니다. 자세한 내용은 AWS Control Tower 설명서의 [제어 참조 가이드](#)를 참조하고 보안 [제어 구현의 사전 예방적 제어](#)를 참조하세요. AWS

제품 수명 주기 관리(PLM)

설계, 개발 및 출시부터 성장 및 성숙도를 거쳐 거부 및 제거에 이르기까지 전체 수명 주기 동안 제품의 데이터 및 프로세스 관리를 나타냅니다.

프로덕션 환경

[환경](#)을 참조하세요.

프로그래밍 가능 로직 컨트롤러(PLC)

제조 분야에서 기계를 모니터링하고 제조 프로세스를 자동화하는 매우 안정적이고 적응력이 뛰어난 컴퓨터입니다.

프롬프트 체이닝

한 [LLM](#) 프롬프트의 출력을 다음 프롬프트의 입력으로 사용하여 더 나은 응답을 생성합니다. 이 기법은 복잡한 작업을 하위 태스크로 나누거나 예비 응답을 반복적으로 세부 조정하거나 확장하는 데 사용됩니다. 이를 통해 모델 응답의 정확성과 관련성을 개선하고 보다 세분화되고 개인화된 결과를 얻을 수 있습니다.

가명화

데이터세트의 개인 식별자를 자리 표시자 값으로 바꾸는 프로세스입니다. 가명화는 개인 정보를 보호하는 데 도움이 될 수 있습니다. 가명화된 데이터는 여전히 개인 데이터로 간주됩니다.

게시/구독(pub/sub)

여러 마이크로서비스에서 비동기 통신을 지원하여 확장성과 응답성을 개선하는 패턴입니다. 예를 들어 마이크로서비스 기반 [MES](#)에서 마이크로서비스는 다른 마이크로서비스가 구독할 수 있는 채널에 이벤트 메시지를 게시할 수 있습니다. 시스템은 게시 서비스를 변경하지 않고도 새 마이크로서비스를 추가할 수 있습니다.

Q

쿼리 계획

SQL 관계형 데이터베이스 시스템의 데이터에 액세스하는 데 사용되는 명령어와 같은 일련의 단계입니다.

쿼리 계획 회귀

데이터베이스 서비스 최적화 프로그램이 데이터베이스 환경을 변경하기 전보다 덜 최적의 계획을 선택하는 경우입니다. 통계, 제한 사항, 환경 설정, 쿼리 파라미터 바인딩 및 데이터베이스 엔진 업데이트의 변경으로 인해 발생할 수 있습니다.

R

RACI 매트릭스

[Responsible, Accountable, Consulted, Informed\(RACI\)](#)를 참조하세요.

RAG

[검색 증강 생성](#)을 참조하세요.

랜섬웨어

결제가 완료될 때까지 컴퓨터 시스템이나 데이터에 대한 액세스를 차단하도록 설계된 악성 소프트웨어입니다.

RASCI 매트릭스

[Responsible, Accountable, Consulted, Informed\(RACI\)](#)를 참조하세요.

RCAC

[행 및 열 액세스 제어](#)를 참조하세요.

읽기 전용 복제본

읽기 전용 용도로 사용되는 데이터베이스의 사본입니다. 쿼리를 읽기 전용 복제본으로 라우팅하여 기본 데이터베이스의 로드를 줄일 수 있습니다.

리아키텍팅

[7R](#)을 참조하세요.

Recovery Point Objective(RPO)

마지막 데이터 복구 시점 이후 허용되는 최대 시간입니다. 이에 따라 마지막 복구 시점과 서비스 중단 사이에 허용되는 데이터 손실로 간주되는 범위가 결정됩니다.

Recovery Time Objective(RTO)

서비스 중단과 서비스 복원 사이의 허용 가능한 지연 시간입니다.

리팩터링

[7R](#)을 참조하세요.

리전

지리적 영역의 AWS 리소스 모음입니다. 각 AWS 리전은 내결함성, 안정성 및 복원력을 제공하기 위해 서로 격리되고 독립적입니다. 자세한 내용은 [계정에서 사용할 수 있는 AWS 리전 지정](#)을 참조하세요.

회귀

숫자 값을 예측하는 ML 기법입니다. 예를 들어, '이 집은 얼마에 팔릴까?'라는 문제를 풀기 위해 ML 모델은 선형 회귀 모델을 사용하여 주택에 대해 알려진 사실(예: 면적)을 기반으로 주택의 매매 가격을 예측할 수 있습니다.

리호스팅

[7R](#)을 참조하세요.

릴리스

배포 프로세스에서 변경 사항을 프로덕션 환경으로 승격시키는 행위입니다.

재배치

[7R](#)을 참조하세요.

리플랫폼

[7R](#)을 참조하세요.

재구매

[7R](#)을 참조하세요.

복원력

중단에 저항하거나 중단을 복구할 수 있는 애플리케이션의 기능입니다. [고가용성](#) 및 [재해 복구](#)는 AWS 클라우드에서 복원력을 계획할 때 일반적인 고려 사항입니다. 자세한 내용은 [AWS 클라우드 복원력](#)을 참조하세요.

리소스 기반 정책

Amazon S3 버킷, 엔드포인트, 암호화 키 등의 리소스에 연결된 정책입니다. 이 유형의 정책은 액세스가 허용된 보안 주체, 지원되는 작업 및 충족해야 하는 기타 조건을 지정합니다.

RACI(Responsible, Accountable, Consulted, Informed) 매트릭스

마이그레이션 활동 및 클라우드 운영에 참여하는 모든 당사자의 역할과 책임을 정의하는 매트릭스입니다. 매트릭스 이름은 매트릭스에 정의된 책임 유형에서 파생됩니다. 실무 담당자 (R), 의사 결정권자 (A), 업무 수행 조언자 (C), 결과 통보 대상자 (I). 지원자는 (S) 선택사항입니다. 지원자를 포함하면 매트릭스를 RASCI 매트릭스라고 하고, 지원자를 제외하면 RACI 매트릭스라고 합니다.

대응 제어

보안 기준에서 벗어나거나 부정적인 이벤트를 해결하도록 설계된 보안 제어입니다. 자세한 내용은 AWS에서 보안 제어 구현의 [대응 제어](#)를 참조하세요.

retain

[7R](#)을 참조하세요.

사용 중지

[7R](#)을 참조하세요.

검색 증강 세대(RAG)

응답을 생성하기 전에 [LLM](#)이 훈련 데이터 소스 외부에 있는 신뢰할 수 있는 데이터 소스를 참조하는 [생성형 AI](#) 기술입니다. 예를 들어 RAG 모델은 조직의 지식 기반 또는 사용자 지정 데이터에 대한 시맨틱 검색을 수행할 수 있습니다. 자세한 내용은 [검색 증강 생성\(RAG\)이란 무엇인가요?](#)를 참조하세요.

교체

공격자가 자격 증명에 액세스하는 것을 더욱 어렵게 만들기 위해 [보안 암호](#)를 주기적으로 업데이트 하는 프로세스입니다.

행 및 열 액세스 제어(RCAC)

액세스 규칙이 정의된 기본적이고 유연한 SQL 표현식을 사용합니다. RCAC는 행 권한과 열 마스크로 구성됩니다.

RPO

[목표 복구 시점\(RPO\)](#)을 참조하세요.

RTO

[목표 복구 시간\(RTO\)](#)을 참조하세요.

런북

특정 작업을 수행하는 데 필요한 일련의 수동 또는 자동 절차입니다. 일반적으로 오류율이 높은 반복 작업이나 절차를 간소화하기 위해 런북을 만듭니다.

S

SAML 2.0

많은 ID 제공업체(idP)에서 사용하는 개방형 표준입니다. 이 기능을 사용하면 연동 SSO(Single Sign-On)를 AWS Management Console 사용할 수 있으므로 사용자는 조직의 모든 사용자에게 대해 IAM에서 사용자를 생성하지 않고도 로그인하거나 AWS API 작업을 호출할 수 있습니다. SAML 2.0 기반 페더레이션에 대한 자세한 내용은 IAM 설명서의 [SAML 2.0 기반 페더레이션 정보](#)를 참조하십시오.

SCADA

[감독 제어 및 데이터 획득](#)을 참조하세요.

SCP

[서비스 제어 정책](#)을 참조하세요.

보안 암호

에는 암호 또는 사용자 자격 증명과 같이 암호화된 형식으로 저장하는 AWS Secrets Manager 키 또는 제한된 정보가 있습니다. 보안 암호 값과 메타데이터로 구성됩니다. 보안 암호 값은 바이너리, 단일 문자열 또는 여러 문자열일 수 있습니다. 자세한 내용은 AWS Secrets Manager 설명서의 [Secrets Manager 보안 암호란 무엇인가요?](#)를 참조하세요.

보안 중심 설계

전체 개발 프로세스에서 보안을 고려하는 시스템 엔지니어링에서의 접근 방식입니다.

보안 제어

위험 행위자가 보안 취약성을 악용하는 능력을 방지, 탐지 또는 감소시키는 기술적 또는 관리적 가드레일입니다. 보안 제어는 [예방](#), [감지](#), [대응](#), [선제적](#)과 같은 기본적인 네 가지 보안 제어 유형으로 구분됩니다.

보안 강화

공격 표면을 줄여 공격에 대한 저항력을 높이는 프로세스입니다. 더 이상 필요하지 않은 리소스 제거, 최소 권한 부여의 보안 모범 사례 구현, 구성 파일의 불필요한 기능 비활성화 등의 작업이 여기에 포함될 수 있습니다.

보안 정보 및 이벤트 관리(SIEM) 시스템

보안 정보 관리(SIM)와 보안 이벤트 관리(SEM) 시스템을 결합하는 도구 및 서비스입니다. SIEM 시스템은 서버, 네트워크, 디바이스 및 기타 소스에서 데이터를 수집, 모니터링 및 분석하여 위협과 보안 침해를 탐지하고 알림을 생성합니다.

보안 응답 자동화

보안 이벤트에 자동으로 응답하거나 이를 해결하도록 설계된 사전 정의되고 프로그래밍된 작업입니다. 이러한 자동화는 보안 모범 사례를 구현하는 데 도움이 되는 [탐지](#) 또는 [대응](#) AWS 보안 제어 역할을 합니다. 자동화된 응답 작업의 예로 VPC 보안 그룹 수정, Amazon EC2 인스턴스 패치 적용 또는 자격 증명 교체 등이 있습니다.

서버 측 암호화

대상에서 데이터를 수신하는 AWS 서비스에 의한 데이터 암호화.

서비스 제어 정책(SCP)

AWS Organizations에 속한 조직의 모든 계정에 대한 권한을 중앙 집중식으로 제어하는 정책입니다. SCP는 관리자가 사용자 또는 역할에 위임할 수 있는 작업에 대해 제한을 설정하거나 가드레일을 정의합니다. SCP를 허용 목록 또는 거부 목록으로 사용하여 허용하거나 금지할 서비스 또는 작업을 지정할 수 있습니다. 자세한 내용은 AWS Organizations 설명서의 [서비스 제어 정책을](#) 참조하세요.

서비스 엔드포인트

에 대한 진입점의 URL입니다 AWS 서비스. 엔드포인트를 사용하여 대상 서비스에 프로그래밍 방식으로 연결할 수 있습니다. 자세한 내용은 AWS 일반 참조의 [AWS 서비스 엔드포인트](#)를 참조하십시오.

서비스 수준에 관한 계약(SLA)

IT 팀이 고객에게 제공하기로 약속한 내용(예: 서비스 가동 시간 및 성능)을 명시한 계약입니다.

서비스 수준 지표(SLI)

오류 발생률, 가용성 또는 처리량과 같은 서비스의 성능 측면에 대한 측정값입니다.

서비스 수준 목표(SLO)

[서비스 수준 지표](#)로 측정되는 서비스의 상태를 나타내는 목표 지표입니다.

공동 책임 모델

클라우드 보안 및 규정 준수를 AWS 위해와 공유하는 책임을 설명하는 모델입니다. AWS 는 클라우드의 보안을 담당하는 반면, 사용자는 클라우드의 보안을 담당합니다. 자세한 내용은 [공동 책임 모델](#)을 참조하십시오.

SIEM

[보안 정보 및 이벤트 관리 시스템](#)을 참조하세요.

단일 장애점(SPOF)

애플리케이션을 중단시킬 수 있는 애플리케이션의 중요한 단일 구성 요소에서 발생하는 장애입니다.

SLA

[서비스 수준 계약](#)을 참조하세요.

SLI

[서비스 수준 지표](#)를 참조하세요.

SLO

[서비스 수준 목표](#)를 참조하세요.

분할 앤 시드 모델

현대화 프로젝트를 확장하고 가속화하기 위한 패턴입니다. 새로운 기능과 제품 릴리스가 정의되면 핵심 팀이 분할되어 새로운 제품 팀이 만들어집니다. 이를 통해 조직의 역량과 서비스 규모를 조정하고, 개발자 생산성을 개선하고, 신속한 혁신을 지원할 수 있습니다. 자세한 내용은 [AWS 클라우드에서 애플리케이션을 현대화하기 위한 단계별 접근 방식](#)을 참조하세요.

SPOF

[단일 장애점](#)을 참조하세요.

스타 스키마

하나의 큰 팩트 테이블을 사용하여 트랜잭션 또는 측정된 데이터를 저장하고 하나 이상의 더 작은 차원 테이블을 사용하여 데이터 속성을 저장하는 데이터베이스 조직 구조입니다. 이 구조는 [데이터 웨어하우스](#)에서 또는 비즈니스 인텔리전스 목적으로 사용하도록 설계되었습니다.

Strangler Fig 패턴

레거시 시스템을 폐기할 수 있을 때까지 시스템 기능을 점진적으로 다시 작성하고 교체하여 모놀리식 시스템을 현대화하기 위한 접근 방식. 이 패턴은 무화과 덩굴이 나무로 자라 결국 속주를 압도하고 대체하는 것과 비슷합니다. [Martin Fowler](#)가 모놀리식 시스템을 다시 작성할 때 위험을 관리하는 방법으로 이 패턴을 도입했습니다. 이 패턴을 적용하는 방법의 예는 [컨테이너 및 Amazon API Gateway를 사용하여 기존의 Microsoft ASP.NET\(ASMX\) 웹 서비스를 점진적으로 현대화하는 방법](#)을 참조하십시오.

서브넷

VPC의 IP 주소 범위입니다. 서브넷은 단일 가용 영역에 상주해야 합니다.

감독 제어 및 데이터 획득(SCADA)

제조 분야에서 하드웨어와 소프트웨어를 사용하여 물리적 자산과 프로덕션 작업을 모니터링하는 시스템입니다.

대칭 암호화

동일한 키를 사용하여 데이터를 암호화하고 복호화하는 암호화 알고리즘입니다.

합성 테스트

사용자 상호 작용을 시뮬레이션하여 잠재적 문제를 감지하거나 성능을 모니터링하는 방식으로 진행되는 시스템 테스트입니다. [Amazon CloudWatch Synthetics](#)를 사용하여 이러한 테스트를 생성할 수 있습니다.

시스템 프롬프트

[LLM](#)에 컨텍스트, 명령 또는 지침을 제공하여 동작을 지시하는 기법입니다. 시스템 프롬프트는 컨텍스트를 설정하고 사용자와의 상호 작용을 위한 규칙을 설정하는 데 도움이 됩니다.

T

tags

AWS 리소스를 구성하기 위한 메타데이터 역할을 하는 키-값 페어입니다. 태그를 사용하면 리소스를 손쉽게 관리, 식별, 정리, 검색, 필터링할 수 있습니다. 자세한 내용은 [AWS 리소스에 태그 지정](#)을 참조하십시오.

대상 변수

지도 ML에서 예측하려는 값으로, 결과 변수라고도 합니다. 예를 들어, 제조 설정에서 대상 변수는 제품 결함일 수 있습니다.

작업 목록

런북을 통해 진행 상황을 추적하는 데 사용되는 도구입니다. 작업 목록에는 런북의 개요와 완료해야 할 일반 작업 목록이 포함되어 있습니다. 각 일반 작업에 대한 예상 소요 시간, 소유자 및 진행 상황이 작업 목록에 포함됩니다.

테스트 환경

[환경](#)을 참조하세요.

훈련

ML 모델이 학습할 수 있는 데이터를 제공하는 것입니다. 훈련 데이터에는 정답이 포함되어야 합니다. 학습 알고리즘은 훈련 데이터에서 대상(예측하려는 답)에 입력 데이터 속성을 매핑하는 패턴을 찾고, 이러한 패턴을 캡처하는 ML 모델을 출력합니다. 그런 다음 ML 모델을 사용하여 대상을 모르는 새 데이터에 대한 예측을 할 수 있습니다.

Transit Gateway

VPC와 온프레미스 네트워크를 상호 연결하는 데 사용할 수 있는 네트워크 전송 허브입니다. 자세한 내용은 AWS Transit Gateway 설명서의 [전송 게이트웨이란 무엇입니까?](#)를 참조하세요.

트렁크 기반 워크플로

개발자가 기능 브랜치에서 로컬로 기능을 구축하고 테스트한 다음 해당 변경 사항을 기본 브랜치에 병합하는 접근 방식입니다. 이후 기본 브랜치는 개발, 프로덕션 이전 및 프로덕션 환경에 순차적으로 구축됩니다.

신뢰할 수 있는 액세스

사용자를 대신하여 AWS Organizations 및 해당 계정에서 조직에서 작업을 수행하도록 지정하는 서비스에 대한 권한 부여. 신뢰할 수 있는 서비스는 필요할 때 각 계정에 서비스 연결 역할을 생성하여 관리 작업을 수행합니다. 자세한 내용은 설명서의 [다른 AWS 서비스와 AWS Organizations 함께 사용](#)을 참조하세요 AWS Organizations .

튜닝

ML 모델의 정확도를 높이기 위해 훈련 프로세스의 측면을 여러 변경하는 것입니다. 예를 들어, 레이블링 세트를 생성하고 레이블을 추가한 다음 다양한 설정에서 이러한 단계를 여러 번 반복하여 모델을 최적화하는 방식으로 ML 모델을 훈련할 수 있습니다.

피자 두 판 팀

피자 두 판이면 충분한 소규모 DevOps 팀. 피자 두 판 팀 규모는 소프트웨어 개발에 있어 가능한 최상의 공동 작업 기회를 보장합니다.

U

불확실성

예측 ML 모델의 신뢰성을 저해할 수 있는 부정확하거나 불완전하거나 알려지지 않은 정보를 나타내는 개념입니다. 불확실성에는 두 가지 유형이 있습니다. 인식론적 불확실성은 제한적이고 불완전한 데이터에 의해 발생하는 반면, 우연한 불확실성은 데이터에 내재된 노이즈와 무작위성에 의해 발생합니다. 자세한 내용은 [Quantifying uncertainty in deep learning systems](#) 가이드를 참조하십시오.

차별화되지 않은 작업

애플리케이션을 만들고 운영하는 데 필요하지만 최종 사용자에게 직접적인 가치를 제공하거나 경쟁 우위를 제공하지 못하는 작업을 헤비 리프팅이라고도 합니다. 차별화되지 않은 작업의 예로는 조달, 유지보수, 용량 계획 등이 있습니다.

상위 환경

[환경](#)을 참조하세요.

V

정리

스토리지를 회수하고 성능을 향상시키기 위해 증분 업데이트 후 정리 작업을 수행하는 데이터베이스 유지 관리 작업입니다.

버전 제어

리포지토리의 소스 코드 변경과 같은 변경 사항을 추적하는 프로세스 및 도구입니다.

VPC 피어링

프라이빗 IP 주소를 사용하여 트래픽을 라우팅할 수 있게 하는 두 VPC 간의 연결입니다. 자세한 내용은 Amazon VPC 설명서의 [VPC 피어링이란?](#)을 참조하십시오.

취약성

시스템 보안을 손상시키는 소프트웨어 또는 하드웨어 결함입니다.

W

웜 캐시

자주 액세스하는 최신 관련 데이터를 포함하는 버퍼 캐시입니다. 버퍼 캐시에서 데이터베이스 인스턴스를 읽을 수 있기 때문에 주 메모리나 디스크에서 읽는 것보다 빠릅니다.

웜 데이터

자주 액세스하지 않는 데이터입니다. 이런 종류의 데이터를 쿼리할 때는 일반적으로 적절히 느린 쿼리가 허용됩니다.

창 함수

현재 레코드와 어떤 식으로든 관련된 행 그룹에서 계산을 수행하는 SQL 함수입니다. 창 함수는 이동 평균을 계산하거나 현재 행의 상대적 위치를 기반으로 행 값에 액세스하는 등의 태스크를 처리하는 데 유용합니다.

워크로드

고객 대면 애플리케이션이나 백엔드 프로세스 같이 비즈니스 가치를 창출하는 리소스 및 코드 모음입니다.

워크스트림

마이그레이션 프로젝트에서 특정 작업 세트를 담당하는 직무 그룹입니다. 각 워크스트림은 독립적이지만 프로젝트의 다른 워크스트림을 지원합니다. 예를 들어, 포트폴리오 워크스트림은 애플리케이션 우선순위 지정, 웨이브 계획, 마이그레이션 메타데이터 수집을 담당합니다. 포트폴리오 워크스트림은 이러한 자산을 마이그레이션 워크스트림에 전달하고, 마이그레이션 워크스트림은 서버와 애플리케이션을 마이그레이션합니다.

WORM

[Write Once, Read Many\(WORM\)](#)를 참조하세요.

WQF

[AWS Workload Qualification Framework](#)를 참조하세요.

Write Once Read Many(WORM)

데이터를 한 번 쓰고 데이터가 삭제되거나 수정되지 않도록 하는 스토리지 모델입니다. 권한 있는 사용자는 필요한 만큼 여러 번 데이터를 읽을 수 있지만 데이터를 변경할 수는 없습니다. 이 데이터 스토리지 인프라는 [변경 불가능](#)한 항목으로 간주됩니다.

Z

제로데이 익스플로잇

[제로데이 취약성](#)을 악용하는 공격(일반적으로 맬웨어)입니다.

제로데이 취약성

프로덕션 시스템의 명백한 결함 또는 취약성입니다. 위협 행위자는 이러한 유형의 취약성을 사용하여 시스템을 공격할 수 있습니다. 개발자는 공격의 결과로 취약성을 인지하는 경우가 많습니다.

제로샷 프롬프팅

태스크를 수행하기 위해 [LLM](#)에 명령을 제공하지만 안내에 도움이 되는 예제(샷)는 제공하지 않습니다. LLM은 사전 훈련된 지식을 사용하여 태스크를 처리해야 합니다. 제로샷 프롬프팅의 효과는 태스크의 복잡성과 프롬프트의 품질에 따라 달라집니다. [퓨샷 프롬프팅](#)도 참조하세요.

좀비 애플리케이션

평균 CPU 및 메모리 사용량이 5% 미만인 애플리케이션입니다. 마이그레이션 프로젝트에서는 이러한 애플리케이션을 사용 중지하는 것이 일반적입니다.

기계 번역으로 제공되는 번역입니다. 제공된 번역과 원본 영어의 내용이 상충하는 경우에는 영어 버전이 우선합니다.