

Happy EduTech 배포 준비성 체크리스트

프로젝트: Happy EduTech AI 수학 학습 플랫폼 배포 일자: 2024년 아키텍처: Lambda + SQS + GPU Auto Scaling (AI/ML 워크로드)

AI/ML 플랫폼 특성상 전통적 CI/CD 대신 MLOps 기반 모델 배포 및 A/B 테스트 검증 방식 적용

사전 배포 검증 (Pre-Deployment)

고객 요구사항 확인

- ☒ 아키텍처 설계 승인: 고객 검토 및 승인 완료
- ☒ 비용 견적 합의: GPU 인스턴스 비용 포함 확인
- ☒ 규정 준수 검토: 교육 데이터 보안 요구사항 충족
- ☒ SLA/RTO/RPO 합의: AI 서비스 가용성 99.5% 목표

인프라 코드 검증

- ☒ CDK 검증: Lambda, SQS, Auto Scaling 스택 검증
- ☒ 개발 환경 배포 테스트: 스테이징 환경 배포 완료
- ☒ 리소스 태깅 표준 적용: 비용 추적 태그 설정
- ☒ IAM 권한 최소화: Lambda, S3, SQS 최소 권한 적용

자동화 테스트

- ☒ 인프라 프로비저닝 테스트: CDK 배포 성공
- ☒ 네트워크 연결성 테스트: VPC, 서브넷 검증
- ☒ 보안 구성 검증: Security Group 규칙 테스트
- ☒ 백업/복구 테스트: S3 버전 관리, 모델 아티팩트 백업

배포 실행 (Deployment)

최종 점검

- ☒ 고객 계정 접근 권한 확인: IAM Role 설정
- ☒ 배포 일정 최종 확인: 학습 서비스 중단 최소화
- ☒ 롤백 계획 준비: 이전 모델 버전 즉시 복원 가능

배포 후 검증 (Post-Deployment)

인프라 검증

- ☒ 리소스 상태 확인: Lambda, SQS, Auto Scaling 정상
- ☒ 모니터링 설정 확인: CloudWatch 알람 + 커스텀 AI 메트릭
- ☒ 고객 접근 권한 확인: 콘솔/CLI 접근 가능
- ☒ 비용 추적 시작: Cost Explorer 설정 완료

AI/ML 배포 방식

MLOps 검증 프로세스

| 단계 | 방법 |
|-------|---------------------|
| 모델 검증 | A/B 테스트 (정확도 비교) |
| 배포 전략 | Blue-Green 배포 |
| 롤백 | 이전 모델 버전 즉시 전환 |
| 모니터링 | OCR 정확도, 추론 지연시간 추적 |

미충족 항목

- ☐ CI/CD 파이프라인 문서화: MLOps 프로세스로 대체
- ☐ 배포 체크리스트 표준화: 모델별 검증 절차로 대체

참고: AI/ML 워크로드 특성상 전통적 소프트웨어 배포 체크리스트 대신 MLOps 검증 프로세스 적용