제131회 컴퓨터시스템응용기술사 해설집

2023.08.26

국가기술자격 기술사 시험문제

기술사 제 131 회 제 3 교시 (시험시간: 100 분)

분야	정보통신	자격종목	컴퓨터시스템응용기술사	수검번호		성명	

※ 다음 문제 중 4 문제를 선택하여 설명하시오. (각 10 점)

1. 현재의 딥러닝 기술은 사람의 눈으로 식별되지 않을 만큼 작은 노이즈를 추가해서 만든 적대적 예제(Adversarial Example)를 활용한 공격에 취약하다. 이와 관련하여 다음을 설명하시오.
   가. White-box 및 Black-box 적대적 공격에 대한 개념과 장단점 비교
   나. 적대적 훈련(Adversarial Training) 및 Defense GAN(Generative Adversarial Networks) 방어기법

2. 개인정보 비식별 처리와 관련하여 다음을 설명하시오.
   가. 개인정보 비식별 처리 유형
   나. 비식별 개인정보의 위험 요인

3. 디스크 여러 개를 활용하여 속도를 높이고 안정성을 향상시키는 기술인 RAID(Redundant Array of Inexpensive Disk) 기술 중 RAID5 와 RAID6 에 대하여 설명하고, 최소 디스크 수량 및 고장 허용 측면에서 비교하여 설명하시오.

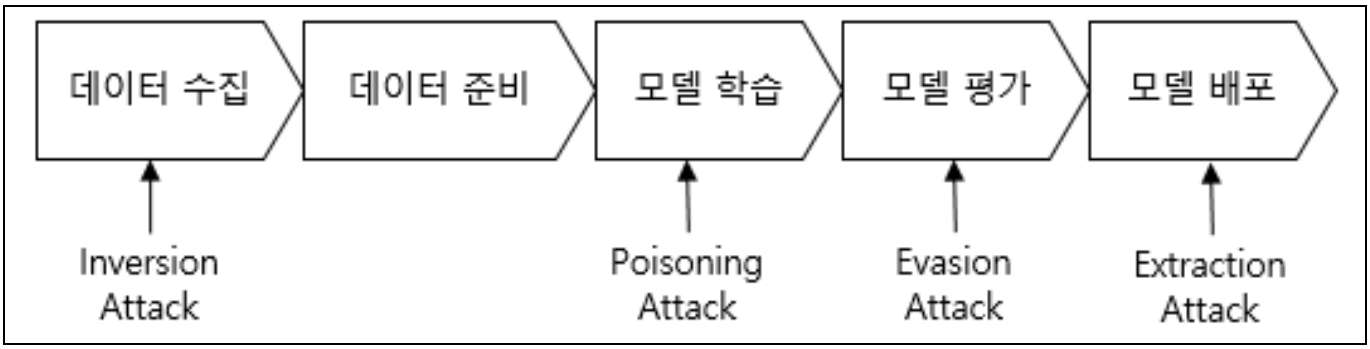
4. 데이터베이스에 사용되는 트랜잭션의 개념과 이를 정의하는 4 가지 중요한 속성을 가리키는 ACID 의 각 요소에 대하여 설명하시오.

5. 공공기관 정보화 사업 추진 시 국가정보원 보안성 검토 절차를 설명하시오.

6. 데이터옵스(DataOps)의 주요 기술을 설명하고, 데브옵스(DevOps)와의 차이점을 설명하시오

01	Adversarial Example		
문제	현재의 딥러닝 기술은 사람의 눈으로 식별되지 않을 만큼 작은 노이즈를 추가해서 만든 적대적 예제(Adversarial Example)를 활용한 공격에 취약하다. 이와 관련하여 다음을 설명하시오. 가. White-box 및 Black-box 적대적 공격에 대한 개념과 장단점 비교 나. 적대적 훈련(Adversarial Training) 및 Defense GAN(Generative Adversarial Networks) 방어기법		
도메인	인공지능	난이도	상(상/중/하)
키워드	노이즈 제거, Defense Model		
출제배경	최근 인공지능 발전에 따른 보안 문제 대두 확인		
참고문헌	https://tutorials.pytorch.kr/beginner/fgsm_tutorial.html		
해설자	전일 기술사(제 114회 정보관리기술사 / rosemachine@naver.com)		

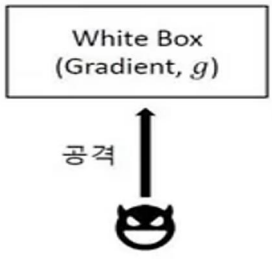
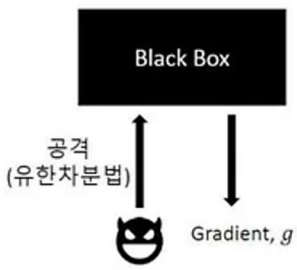
I. 머신 러닝 알고리즘 취약점, AI 적대적 공격의 개요



- 머신 러닝 진행 단계별 취약점을 공격하여 머신 러닝 학습 결과에 오류를 발생시키거나 모델 추출을 통한 모델의 악의적 사용을 위한 머신 러닝 알고리즘 공격 기법

II. White-box 및 Black-box 적대적 공격에 대한 개념과 장단점 비교

가. White-box 및 Black-box 적대적 공격 개념

구분	개념도	개념
White-box 적대적 공격		- 공격자가 모델에 대해 아키텍처, 입력, 출력, 가중치를 포함한 모든 것을 알고 있고 접근할 수 있다고 가정한 공격
Black-box 적대적 공격		- 공격자가 모델의 입력과 출력에 대해서만 접근 가능하고 모델의 가중치와 아키텍처에 관한 내용은 모른다고 가정한 공격

나. White-box 및 Black-box 적대적 공격 장단점 비교

구분	White-box 적대적 공격	Black-box 적대적 공격
장점	- 정교한 공격 가능 - 알고리즘 및 기법에 대한 전반적 이해 높아 예외 상황 발생 시 즉각적 대응 가능 - 공격 기법의 다양성 및 확장 가능 우수	- 단기간 공격 가능 - 자원 소요 상대적 불필요
단점	- 학습 기간 장기화 - 자원 투입 多 - 공격 기법의 복잡성으로 인해 숙련도에 따른 성공률 높음	- 공격 효율성 낮음

- White-box 및 Black-box 적대적 공격에 대응 가능한 적대적 훈련 기법 존재

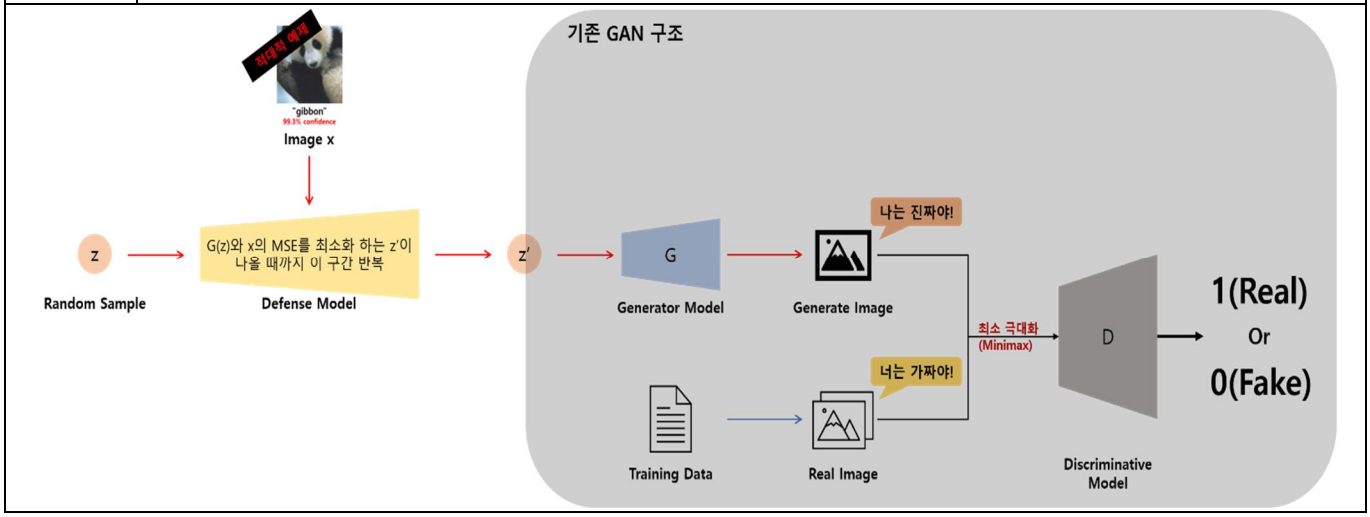
III. 적대적 훈련 및 Defense GAN 방어 기법

가 적대적 훈련(Adversarial Training) 상세 설명

구분	세부 내용	
정의	- 모델을 학습시킬 때 적대적 공격에 대응하기 위해 기존 학습 데이터 셋에 적대적 샘플들을 추가로 학습시켜 적대적 공격에 대해 강인하게 만드는 기법	
특징	적대적 방어 기술	- White Box 적대적 공격과 Black box 적대적 공격 대응 가능
	간편성	- 이미 만들어진 알고리즘을 재활용
	학습효과 우수	- 훈련 시간 및 학습량에 비례하여 적대적 공격 방어 우수

나. Defense GAN(Generative Adversarial Networks) 방어 기법 상세 설명

구분	세부 내용
정의	- 모델을 학습시킬 때 적대적 공격에 대응하기 위해 기존 학습 데이터 셋에 적대적 샘플들을 추가로 학습시켜 적대적 공격에 대해 강인하게 만드는 기법



원리	- Defense-GAN은 생성 이미지와 적대적 예제의 차이를 최소화하는 새로운 생성 데이터(z')을 생성 - 새롭게 만들어진 생성 데이터는 기존 GAN 학습이 이루어짐 - 새로운 생성 데이터(z')를 통해 GAN을 실시하게 되면 자연스럽게 적대적 예제의 노이즈는 제거되는 효과를 가져옴

- Defense-GAN이 각광받는 이유는 기존 모델의 분류기(Classifier)의 수정 없이 공격에 대한 방어가 가능

"끝"

02	개인정보 비식별 처리		
문제	개인정보 비식별 처리와 관련하여 다음을 설명하시오. 가. 개인정보 비식별 처리 유형 나. 비식별 개인정보의 위험 요인		
도메인	보안	난이도	중(상/중/하)
키워드	가명화, 총계처리, 데이터범주화, 데이터삭제, 데이터 마스킹		
출제배경	개인정보 비식별 처리에 대한 이해 확인		
참고문헌	ITPE 기술사회 자료집		
해설자	정상반멘토 이상헌 기술사(제 118회 정보관리기술사 / bluesanta97@naver.com)		

I. 개인식별 요소 제거, 개인정보 비식별 처리 개요

정의	- 개인정보의 일부 혹은 전부를 삭제하거나 변형을 통해 특정 개인을 식별할 수 없도록 취하는 조치

II. 개인정보 비식별 처리 유형

가. 개인정보 비식별 처리 유형 설명

처리 유형	설명	대상
가명화	- 개인정보 중 주요 식별요소를 다른 값으로 대체하는 기법	- 성명, 기타 고유특징
총계처리	- 총계 값을 적용하여 개인을 식별할 수 없게 하는 조치	- 개인과 직접 관련된 날짜 정보 - 기타 고유 특징
데이터삭제	- 개인 식별이 가능한 데이터 삭제 처리	- 개인을 식별할 수 있는 정보 - 생체정보

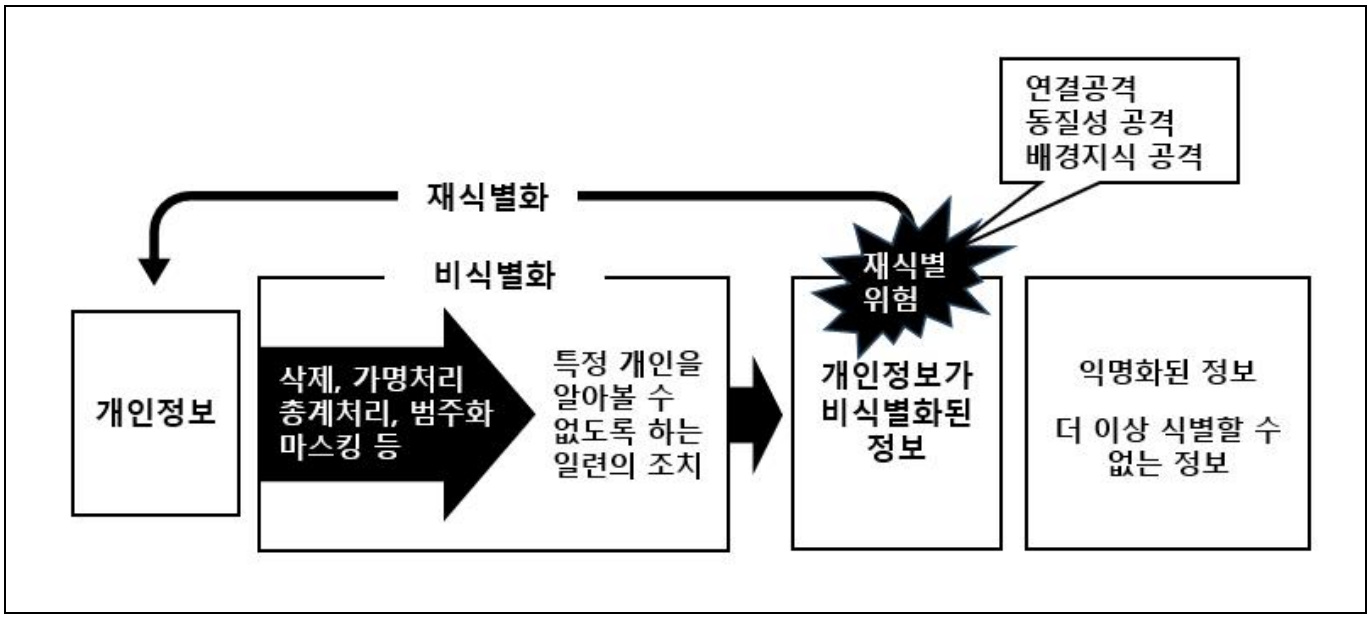
데이터 범주화	- 특정 정보를 대표 값으로 변환, 구간 값으로 변환하는 조치	- 개인을 식별할 수 있는 정보 - 기관·단체 등의 이용자 계정
데이터 마스킹	- 데이터의 전부 또는 일부분을 대체값(공백, 노이즈 등)으로 변환	- 쉽게 개인을 식별할 수 있는 정보 - 기관·단체 등의 이용자 계정

나. 개인정보 비식별 처리 유형 별 세부 기술 설명

처리 유형	세부기술	설명
가명화	휴리스틱 가명화	- 식별자에 해당하는 값들을 몇 가지 정해진 규칙으로 대체하거나 사람의 판단에 따라 가공하여 자세한 개인정보를 숨기는 방법
	암호화	- 정보 가공 시 일정한 규칙의 알고리즘을 적용하여 암호화함으로써 개인정보를 대체하는 방법
	교환 방법	- 기존의 데이터베이스의 레코드를 사전에 정해진 외부의 변수(항목)값과 연계하여 교환
총계처리	총계처리	- 데이터 전체 또는 부분을 집계(총합, 평균 등)
	부분 총계	- 데이터 셋 내 일정부분 레코드만 총계 처리함. 즉, 다른 데이터 값에 비하여 오차 범위가 큰 항목을 통계값(평균 등)으로 변환
	라운딩	- 집계 처리된 값에 대하여 라운딩(올림, 내림, 사사오입) 기준을 적용하여 최종 집계 처리하는 방법
	재배열	- 기존 정보값은 유지하면서 개인이 식별되지 않도록 데이터를 재배열하는 방법
데이터 삭제	식별자 삭제	- 원본 데이터에서 식별자를 단순 삭제하는 방법
	식별자 부분 삭제	- 식별자 전체를 삭제하는 방식이 아니라, 해당 식별자의 일부를 삭제하는 방법
	레코드 삭제	- 다른 정보와 뚜렷하게 구별되는 레코드 전체를 삭제하는 방법
	식별요소 전부삭제	- 식별자뿐만 아니라 잠재적으로 개인을 식별할 수 있는 속성자까지 전부 삭제하여 프라이버시 침해 위험을 줄이는 방법
데이터 범주화	감추기	- 명확한 값을 숨기기 위하여 데이터의 평균 또는 범주 값으로 변환하는 방식
	랜덤 라운딩	- 수치 데이터를 임의의 수 기준으로 올림(round up) 또는 내림(round down)하는 기법
	범위 방법	- 수치데이터를 임의의 수 기준의 범위(range)로 설정하는 기법
	제어 라운딩	- 랜덤 라운딩 방법에서 어떠한 특정 값을 변경할 경우 행과 열의 합이 일치하지 않는 단점 해결을 위해 행과 열이 맞지 않는 것을 제어하여 일치시키는 기법
데이터 마스킹	임의 잡음 추가	- 개인 식별이 가능한 정보에 임의의 숫자 등 잡음을 추가(더하기 또는 곱하기)하는 방법
	공백과 대체	- 특정 항목의 일부 또는 전부를 공백 또는 대체문자(' * ', ' _ ' 등이나 전각 기호)로 바꾸는 기법

III. 비식별 개인정보의 위험 요인

가. 비식별 개인정보의 위험 요인



- 개인정보 비식별 처리 이후에도 비식별 개인정보에 대해 5가지 위험요인이 존재함.

나. 비식별 개인정보의 위험 요인 상세 설명

위험요인	설명	보호모델
연결공격	- 비식별 조치된 결과와 다른 공개 데이터간 결합을 통해 개인을 식별하는 공격	k-익명성
동질성 공격	- 범주화 된 K-익명성 데이터 집합에서 동일한 정보를 이용하여 대상의 정보를 알아내는 공격	l-다양성
배경지식 공격	- 공격자의 배경 지식을 통해 대상의 민감정보를 알아내는 공격	
쏠림 공격	- 정보가 특정한 값에 쏠려 있을 경우 확률적으로 민감 정보를 추론할 수 있는 공격	t-근접성
유사성 공격	- 비식별 조치된 정보가 서로 다르지만 의미상 유사하다면 민감정보를 유추할 수 있는 공격	

- 여러 비식별 개인정보에 대한 위험요인은 프라이버시 보호모델을 통하여 재식별을 방지하여야 함

IV. 비식별화 조치 이후 재식별 방지 프라이버시 보호모델. k-익명성, l-다양성, t-근접성

유형	설명	적용 사례
k-익명성	- 특정인임을 추론할 수 있는지 여부를 검토, 일정 확률 수준 이상 비식별 되도록 하는 기법	- 동일한 값을 가진 레코드를 k개 이상으로 함 - 이 경우 특정 개인을 식별할 확률 : 1/k
l-다양성	- 특정인 추론이 안된다고 해도 민감한 정보의 다양성을 높여 추론 가능성을 낮추는 기법	- 각 레코드는 최소 l개 이상의 다양성을 가지도록 하여 동질성 또는 배경 지식 등에 의한 추론 방지
t-근접성	- l-다양성뿐만 아니라, 민감한 정보의 분포를 낮추어 추론 가능성을 더욱 낮추는 기법	- 전체 데이터 집합의 정보 분포와 특정 정보의 분포 차이를 t 이하로 함

"끝"

03	RAID(Redundant Array of Inexpensive Disk)		
문제	디스크 여러 개를 활용하여 속도를 높이고 안정성을 향상시키는 기술인 RAID(Redundant Array of Inexpensive Disk) 기술 중 RAID5와 RAID6에 대하여 설명하고, 최소 디스크 수량 및 고장 허용 측면에서 비교하여 설명하시오.		
도메인	CA/OS	난이도	하(상/중/하)
키워드	미러링, 스트라이핑, 패리티, 저렴한 비용		
출제배경	데이터 백업 및 활용에 대한 이슈가 있어, 관련 지식 확인		
참고문헌	ITPE 기술사회 자료		
해설자	모멘텀 안수현 기술사(제 119회 정보관리기술사 / itpe.momentum@gmail.com)		

I. 시스템의 고가용성과 성능 향상을 위한 스토리지 솔루션, RAID의 개요

가. RAID(Redundant Array of Independent Disks)의 정의

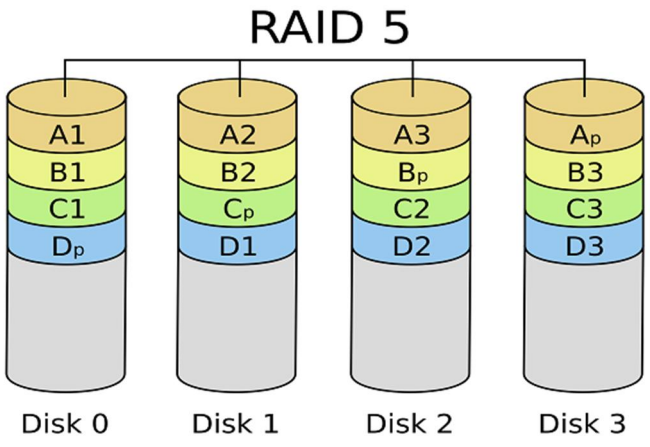
- 여러 개의 디스크에 데이터 중복 저장과 데이터의 분산 저장을 통한 성능 향상과 안정성 향상을 위한 스토리지 솔루션

나. RAID의 유형

유형	설명
RAID 0	Striping 기반의 병렬 디스크 구성 기법
RAID 1	동일하게 복사한 미러 디스크로 구성된 RAID
RAID 5	데이터와 패리티가 분산되어 저장된 구조의 RAID
RAID 6	RAID5 구조와 유사하지만 Double 패리티 사용하는 RAID

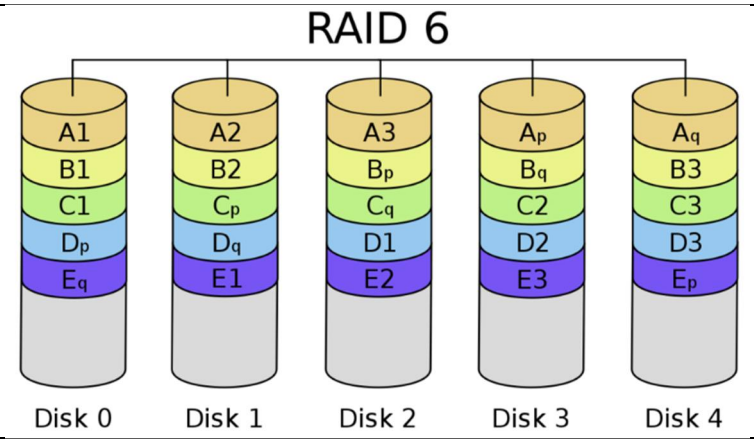
- RAID 2~4도 존재하지만, 전용 Parity 디스크 개념으로 거의 사용하지 않음

II. RAID 5의 설명

구성	
개념	데이터를 블록 단위로 분산하고 패리티도 분산하여 통합 저장한 하드디스크로 구성된 RAID로, 블록 인터리브된 분산 패리티(Block-interleaved striping with distributed parity) 방식이라고도 함

장단점	장점	단점
	- 처리량과 성능이 좋음 - 데이터 중복성을 제공 - 실패한 드라이브의 신속한 재구축 - 사용 가능한 공간을 효율적으로 사용. - 가장 널리 사용되는 RAID 구현 중 하나	- 패리티로 인해 다른 구성에 비해 쓰기 속도가 느림 - 데이터 복원에는 시간 소요 - 구현이 복잡

III. RAID 6의 설명

구성		
개념	RAID 5와 유사한 구조로, Block 레벨의 Striping과 Double Parity 사용(Parity 분산 제공)하는 RAID		
장단점	장점	단점	
	- 높은 중복성과 내결함성. - 합리적으로 빠른 읽기 작업. - 데이터 접근성이 뛰어남	- 이중 패리티로 인해 쓰기 트랜잭션이 매우 느림 - 구현이 복잡 - 복원에 시간이 많이 걸림	

IV. 최소 디스크 수량 및 고장 허용 측면에서 RAID5와 RAID6의 비교

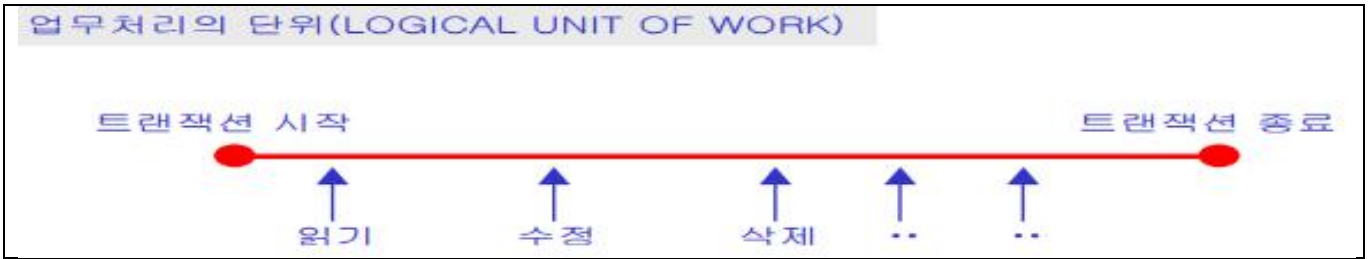
항목	RAID 5	RAID 6
최소 수량	3개	4 개
고장 허용	1 디스크	2 디스크
공간 효율	1 - 1/n (n : disk 수량)	1 - 2/n (n : disk 수량)
읽기 향상	n-1 배	해당 없음
Parity	1	2

"끝"

04	DB 트랜잭션		
문제	데이터베이스에 사용되는 트랜잭션의 개념과 이를 정의하는 4가지 중요한 속성을 가리키는 ACID의 각 요소에 대하여 설명하시오.		
도메인	데애터베이스	난이도	하(상/중/하)
키워드	논리적 작업단위, 원자성, 일관성, 격리성, 영속성		
출제배경	데이터베이스 트랜잭션에 대한 개념 숙지 확인		
참고문헌	ITPE 기술사회 자료집		
해설자	이제원 기술사(제130회 정보관리기술사 / bwmslove@naver.com)		

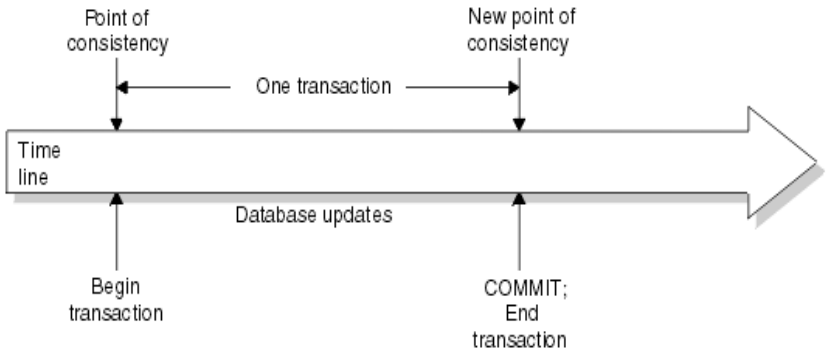
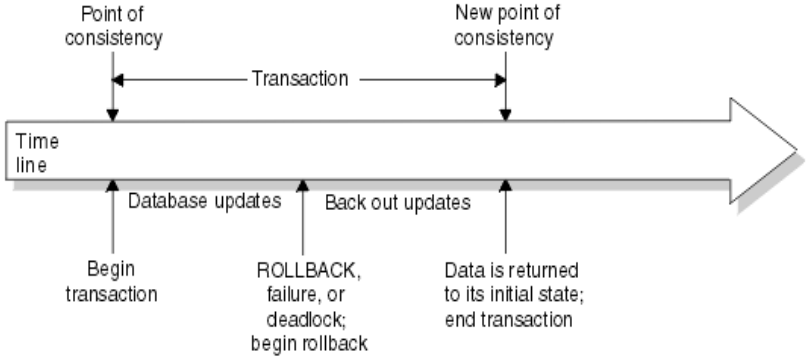
I. DB 논리적 잡업단위, DB 트랜잭션의 개념 및 연산

가. 트랜잭션의 개념



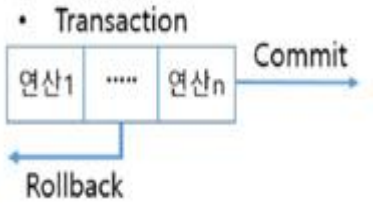
개념	한번에 처리되어야 할 하나 또는 둘 이상의 일련의 작업단위로써 데이터베이스에 행해지는 작업의 논리적 단위

나. 트랜잭션의 연산

구분	개념도	설명
Commit 연산		- 트랜잭션 성공 수행 되었을 시 수행 - 일관된 상태 유지 - Commit 연산의 실행을 통해 최종 DB 반영
Rollback 연산		- 트랜젝션 수행 실패시 선언 - 트랜잭션 수행되기 전의 상태로 돌아감 - 트랜젝션 수행 전의 일관된 상태로 되돌림

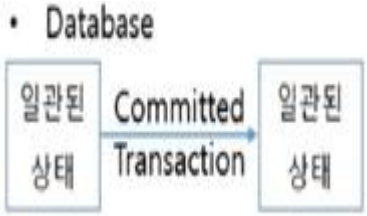
II. 트랜잭션 ACID의 원자성 및 일관성

가. 트랜잭션 ACID의 원자성

특징	개념도	설명
원자성 (Automicity)		- 분해가 불가능한 수행 단위로 완전히 수행, 수행되지 않음 - **구현기법: All or Nothing, Commit or Rollback** - **기능: 회복**

- 중간 단계에서 실패하면 모든 변경 사항 롤백

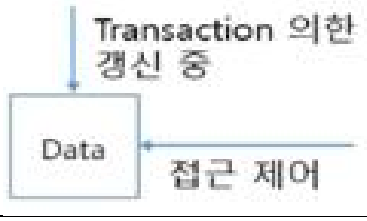
나.트랜젝션 ACID의 일관성

특징	개념도	설명
일관성 (Consistency)		- 트랜잭션이 성공적으로 완료되면 언제나 모순이 없는 상태 유지 - 무결성 제약조건, 사용자가 요구하는 논리적 요건의 충족 - **구현기법: 도메인 무결성, 릴레이션 무결성** - **기능: 무결성 제약조건, 동시성 제어**

- 트랜젝션이 수행되기 전과 수행된 후의 데이터베이스 상태의 일관성

III. 트랜잭션 ACID의 격리성 및 영속성

가. 트랜잭션 ACID의 격리성

특징	개념도	설명
격리성 (Isolation)		- 트랜잭션 실행 중에는 다른 트랜잭션 접근 불가 - **구현기법: isolation level** - **기능: 동시성 제어**

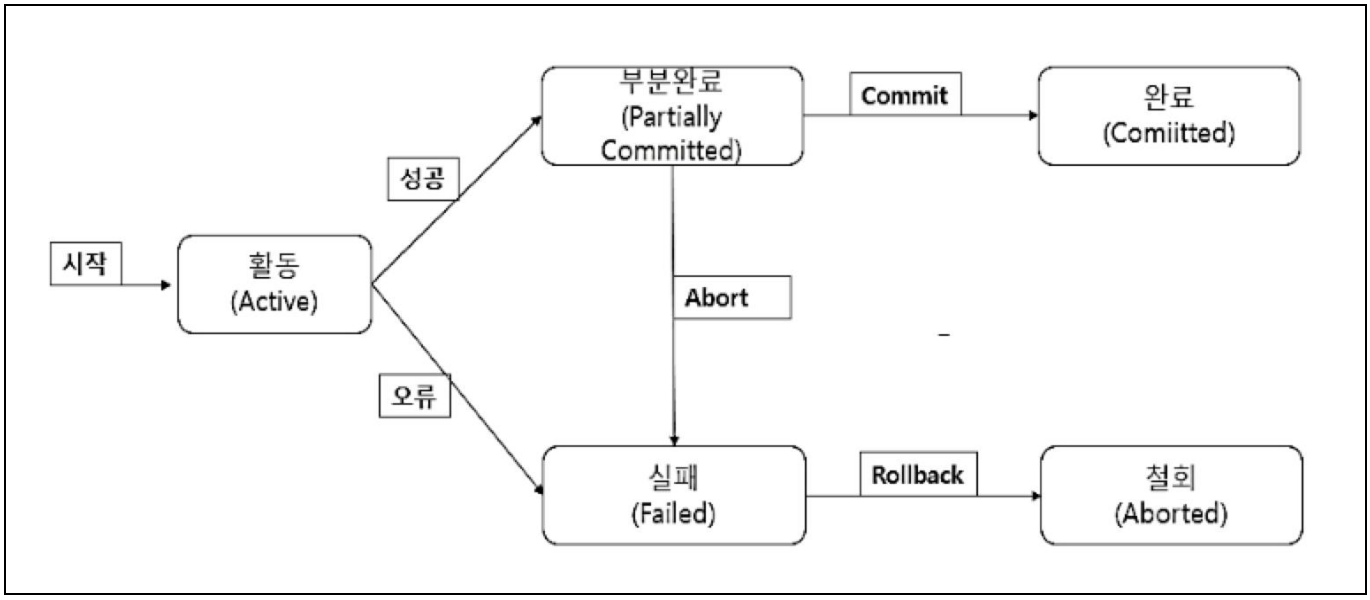
- 각 트랜잭션 간에 서로 영향을 미치지 않는 특징

나. 트랜젝션 ACID의 일관성

특징	개념도	설명
영속성 (Durability)		- 트랜잭션이 실행이 성공적으로 완료하면 그 결과는 영속적임 - **구현기법: Archive, 로그, Redo/Undo 기반 회복** - **기능: 회복**

- 시스템 장애 또는 전원 공급 장애와 같은 문제가 발생해도 트랜잭션 결과는 유지되어야 하는 특징

IV. 데이터베이스 트랜잭션 상태 전이도

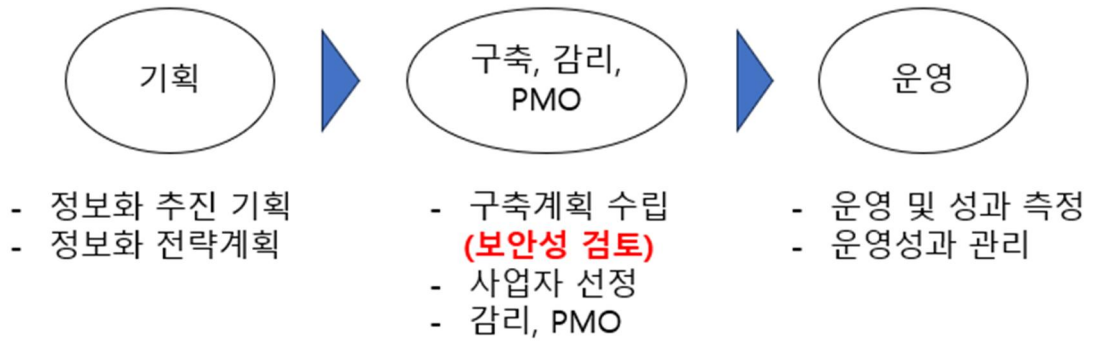


- 다중 프로그램 환경에서 트랜잭션의 ACID 조건을 만족시키기 위해서는 트랜잭션이 동시에 수행될 경우 발생되는 문제점을 해결 위한 동시성 제어 필요

"끝"

05	보안성 검토		
문제	공공기관 정보화 사업 추진 시 국가정보원 보안성 검토 절차를 설명하시오.		
도메인	보안	도메인	보안
키워드	기획, 구축, 운영, 구축계획 수립단계, 범위설정, 보안취약점		
출제배경	공공정보화사업 단계별 사업관리가이드 내 보안성검토 관련 내용 이해하고 있는지 확인		
참고문헌	정보보안 안전성 확보하는 보안성 검토(http://www.itdaily.kr/news/articleView.html?idxno=91114)		
해설자	정유나 기술사(제 130회 정보관리기술사 / audfla89@naver.com)		

I. 정보보안 안전성 확보 활동, 국가정보원 보안성 검토 개요



- 공공기관에서 정보시스템, SW 등을 도입하고 개발하는 사업을 진행할 때 보안 취약점을 사전에 제거하고 체계적인 보안 관리를 하기 위해 수행하는 일련의 과정

II. 국가정보원 보안성 검토 대상 정보화 사업

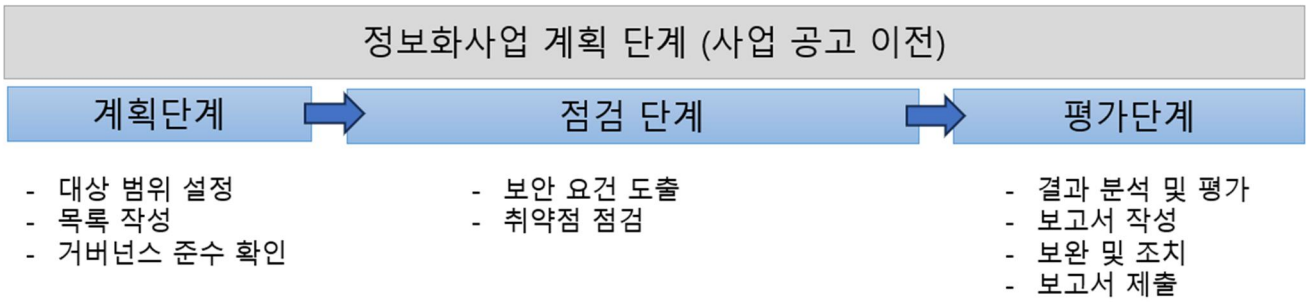
대상 사업	설명
비밀·국가안보·정부정책 관련 사업	비밀 등 국가기밀의 유통·관리와 관련된 정보시스템 구축
	외교·국방 등 국가안보상 중요한 정보통신망 구축
	재난 대비 등 국가위기관리와 관련된 정보통신망 구축
대규모 예산투입과 다량의 DB자료를 처리하는 사업	대규모 정보시스템 구축(10억 원 이상) 정보화 사업
	다량의 개인정보(100만 명 이상)를 처리하는 정보시스템 구축
	지리・환경 정보 등 국가차원의 통합DB 구축
외부기관간 망연동 등 보안상 취약 사업	다수 기관이 공동 활용하기 위한 정보시스템 구축 및 망 연동
	내부 전산망 또는 폐쇄망을 인터넷이나 타 기관의 전산망 등 다른 정보통신망과 연동하는 경우
	원격근무 지원시스템 구축
보안정책 과제 및 최신 IT기술 적용	업무망·인터넷 분리 정보화 사업
	업무망 ↔ 인터넷간 자료교환 시스템 구축
	스마트폰 등 첨단 IT기술을 업무에 활용하는 시스템 구축

기타	국가정보원장 및 중앙행정기관, 광역자치단체장이 보안성검토가 필요하다고 판단하는 정보화사업

- 보안성 검토는 기획, 설계, 검수단계로 구분하여 수행되며 사업 계획단계(사업 공고 전)에 이행해야 함

III. 국가정보원 보안성 검토 절차 개요

가. 국가정보원 보안성 검토 절차 개념도



- 서버, NW, 보안장비 취약점 점검 수행 후 보안성 검토 결과를 상급기관에 제출

나. 국가정보원 보안성 검토 절차 상세 설명

단계	설명	
계획 단계	보안성 검토 범위 설정	보안성 검토를 수행할 서버, NW, 보안장비, 응용프로그램의 범위와 목표 정의
	목록 작성	검토 대상인 서버, NW, 보안장비, 응용프로그램을 식별하고 목록 작성
	리스크 평가 및 거버넌스 준수 확인	검토대상의 중요도와 위험 평가하고 정보보안 관련 법륜 준수 여부 검토
점검 단계	보안 요건 도출	보안 점검절차와 방법을 정의하는 보안 요건 도출
	취약점 점검	정의한 계획에 따라 취약점 스캔, 코드 리뷰 등을 이용하여 보안취약점 탐지
평가 단계	결과 분석 및 평가	검토 결과 분석하여 발견된 취약점에 대한 분석 수행
	보고서 작성	취약점, 위험평가, 추진 사항, 개선방안을 상세히 기술
	보완 및 조치	발견된 취약점에 대해 보안 패치, 업데이트, 보안정책 변경 수행
	보고서 제출 및 평가	완료된 검토 보고서를 상급기관 또는 국가정보원에 검토 및 승인 요청

- 각 공공기관은 상급기관의 검토결과를 통보받은 경우 검토결과를 반영하여 보안대책을 보완해야 함

"끝"

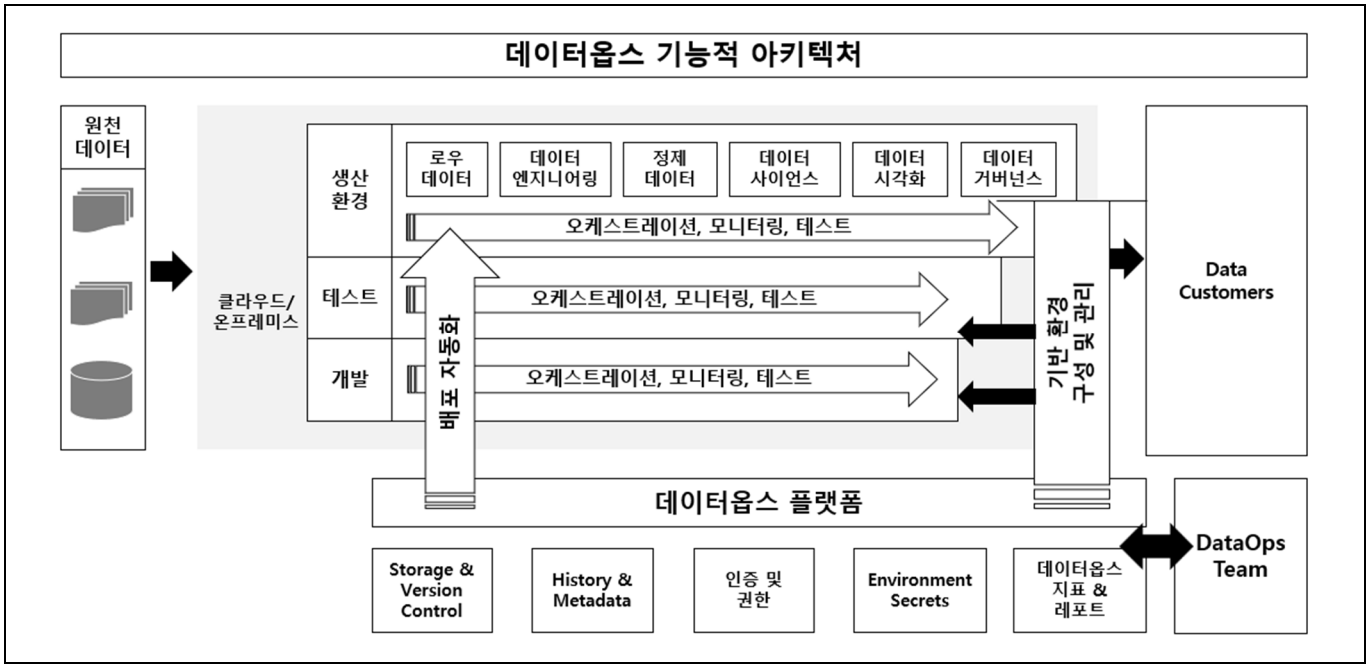
06	데이터옵스(DataOps)		
문제	데이터옵스(DataOps)의 주요 기술을 설명하고, 데브옵스(DevOps)와의 차이점을 설명하시오		
도메인	데이터베이스	난이도	중 (상/중/하)
키워드	샌드박스, 스테이징, 프로덕션, 가치 파이프라인, 혁신 파이프라인, 애자일, 데브옵스, 통계적 제어 프로세스, CI/CD		
출제배경	130회 정보관리 데이터옵스 교차 출제		
참고문헌	ITPE 기술사회 자료		
해설자	정상반멘토 이상헌 기술사(제 118회 정보관리기술사 / bluesanta97@naver.com)		

I. 기업 데이터의 깊이있는 인사이트 획득 기법, 데이터옵스(DataOps)의 개요

정의	- 조직 전체의 데이터 관리자와 데이터 소비자 간의 데이터 흐름의 커뮤니케이션, 통합 및 자동화를 개선하는 데 중점을 둔 협업 데이터 관리 방식
특징	효율적인 데이터 플로우	- 전체 프로세스를 관리하며 데이터 수명주기의 모든 단계에서 연결 및 최적화된 데이터를 선별, 통제, 관리
	안전하고 규정을 준수하는 데이터	- 자동화되고 사용자 지정 가능한 데이터 품질, 마스킹, 토큰화 등에 대한 제어를 적용하여 데이터를 보호하고 여정의 모든 단계에서 규정 준수를 확인
	데이터 비용 절감	- IT 의존도를 줄이고 분석 결과를 가속화하며 데이터 비용을 낮추는 동시에 데이터를 쉽게 검색, 선택 및 프로비저닝 할 수 있음

II. 데이터옵스(DataOps) 주요 기술

가. 데이터옵스(DataOps) 주요 기술 아키텍처



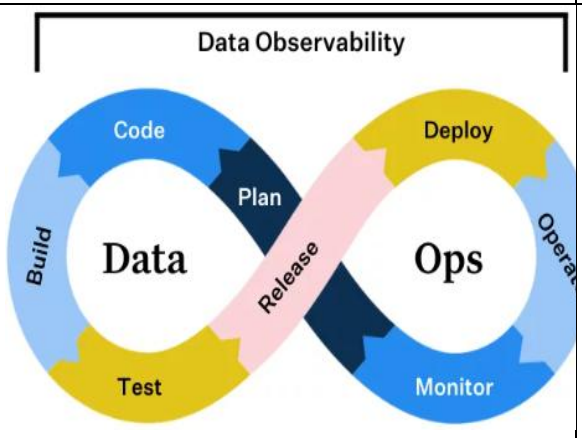
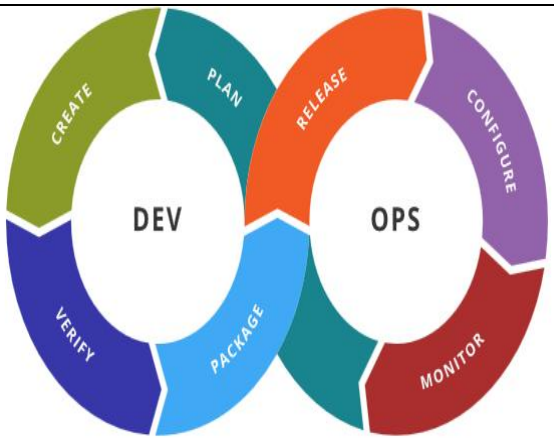
나. 데이터옵스(DataOps) 주요 기술 상세설명

주요 기술	설명
스토리지/리비전 제어	- 버전 제어는 인위적인 변경 사항을 관리. 거버넌스 및 반복 개발에 필수 (예, git, dockerhub)
이력 및 메타 데이터	- 시스템 및 활동 로그 관리 (예, MongoDB)
인증 및 권한	- 환경에 대한 액세스 제어 (예, Auth0)
환경 비밀	- 환경 내 도구 및 리소스에 대한 역할 기반 액세스 (예, Vault)
데이터옵스 지표 및 보고서	- 분석 및 데이터 팀의 상태에 평가에 대한 내부 분석 : CDO 대시 보드 (예, Tableau)
자동 배포	- 하나의 환경에서 프로덕션 환경으로 코드/구성을 이동하는 과정 (예, Jenkins, CircleCI)
환경 생성 및 관리	- 하드웨어, 소프트웨어, 테스트 데이터 세트 등 필요한 모든 것을 가지고 작업할 수 있는 환경을 생성할 수 있는 코드와 같은 인프라 취급 (예, Chef, Puppet)
오케스트레이션, 테스트, 모니터링	- 파이프라인이 실행되는 동안 관련된 모든 도구를 오케스트레이션하고 테스트 및 모니터링하며 문제가 발생 시 경고(예: Airflow, Great Expectations, Grafana).

- 데이터옵스는 데브옵스 접근법을 이용하여 데이터 관리를 수행하나 둘 간에는 명확한 차이가 있음

III. 데이터옵스(DataOps)와 데브옵스(DevOps) 차이점 설명

가. 데이터옵스와 데브옵스의 개념적 차이점

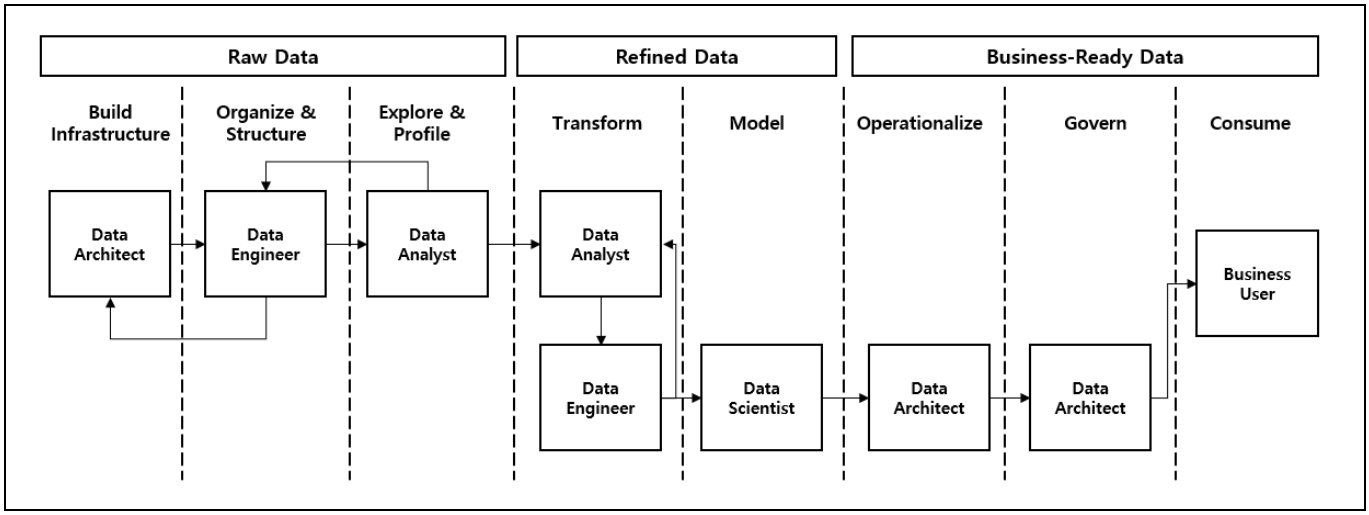
구분	데이터옵스 (DataOps)	데브옵스 (DevOps)
개념	- 조직 전체의 데이터 관리자와 데이터 소비자 간의 데이터 흐름의 커뮤니케이션, 통합 및 자동화를 개선하는 데 중점을 둔 협업 데이터 관리 방식 - 데이터를 분석해 애플리케이션을 형성한 후 최종 사용자에게 신뢰할 수 있는 고품질 데이터를 신속히 제공하기 위한 기본적인 데이터 운영/관리 방식	- 시스템 개발자와 운영을 담당하는 정보기술 전문가 사이의 소통, 협업, 통합 및 자동화를 강조하는 소프트웨어 개발론 - 소프트웨어 제품이나 서비스를 알맞은 시기에 출시하기 위해서 개발과 운영이 상호의존 대응 - 개발과 운영의 합성어 - 개발과 운영의 원활한 상호 작용을 하게 하는 모든 개발 방법론
프레임워크		

나. 데이터옵스와 데브옵스의 세부 차이점 설명

구분	데이터옵스	데브옵스
목적	- 데이터 파이프라인 자동화/모니터링 관리	- SW 개발 자동화 및 모니터링
조합	- 데이터 공학, 데이터 통합,	- SW 개발, 품질보증(QA), 기술운영
협력	- 데이터 품질, 데이터 보안, 개인정보	- SW 엔지니어(개발자),
기대효과	- 데이터 엔지니어, 데이터 과학자,	- 시스템관리자(운영 팀), 테스트

- 데이터옵스를 적용하기 위해서는 기술환경과 분석 전문가, 데이터 엔지니어 등의 역향 확보 필요

IV. 데이터옵스를 활용한 데이터 워크플로우 사례



- 초기 원천 데이터에서 비즈니스에 활용 가능한 데이터까지 데이터옵스 기반 워크플로우를 구성하여 데이터 거버넌스 실행

"끝"

ITPE 기술사회
제131회 정보처리기술사 기출문제 해설집

대 상 정보관리기술사, 컴퓨터시스템응용기술사, 정보통신기술사, 정보시스템감리사 시험
발행일 2023년 08월 26일
집 필 강정배PE, 전일PE, 이상헌PE, 안수현PE, 이제원PE, 정유나PE
출 판 **ITPE(Information Technology Professional Engineer)**
주 소 ITPE 대치점 서울시 강남구 선릉로 86길 17 선릉엠티빌딩 7층
ITPE 선릉점 서울시 강남구 선릉로 86길 15, 3층 IT교육센터 아이티피이
ITPE 강남점 서울시 강남구 테헤란로 52길 21 파라다이스벤처타워 3층 303호
ITPE 영등포점 서울시 영등포구 당산동2가 하나비즈타워 7층 ITPE
연락처 070-4077-1267 / itpe@itpe.co.kr