

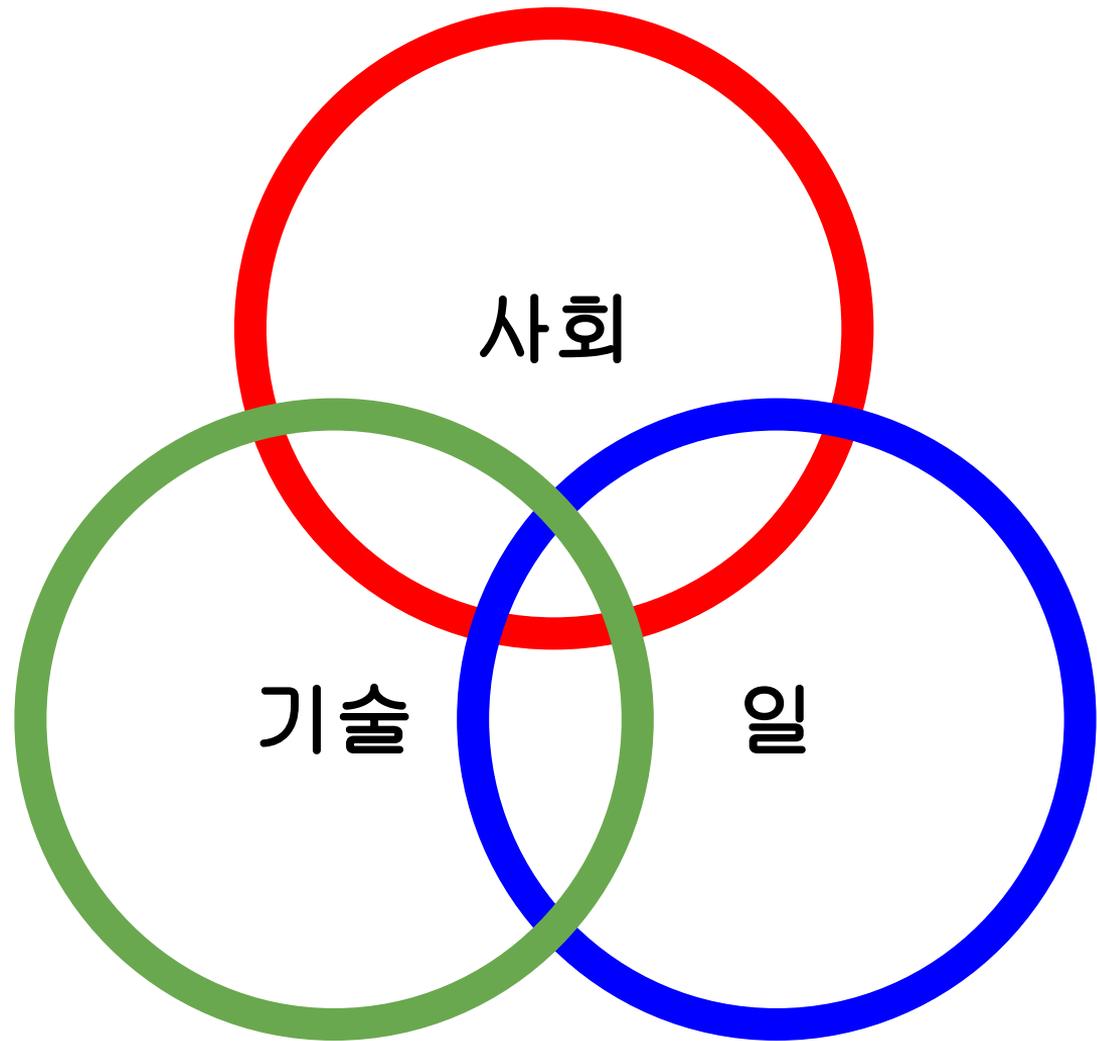
2019년 제3차 전자기록관리 연구포럼 (190731)

기록시스템과 전자기록 장기보존의 이해

외부강사 : 이 연 창 (한국문헌정보기술(주) 아카이브연구소 소장)

openchang.lee@kait.co.kr

변화(變化)



기록 방식의 변화(變化)



0110100010
1101010010
1010101011
01100000....

기록 시스템

시스템

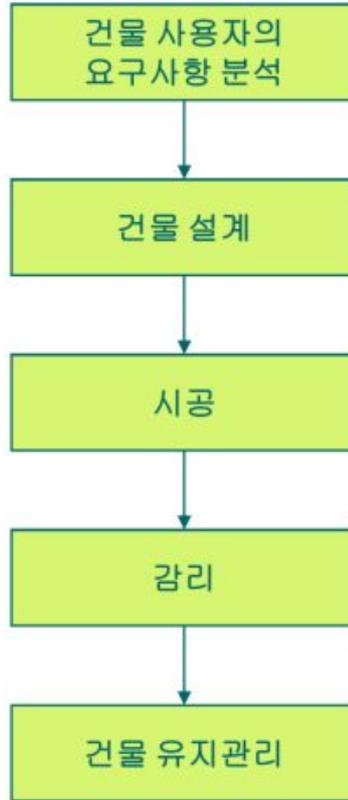
System 體系



〈그림 11-1〉 시스템의 3층 구조

SW 개발과정

건축



소프트웨어

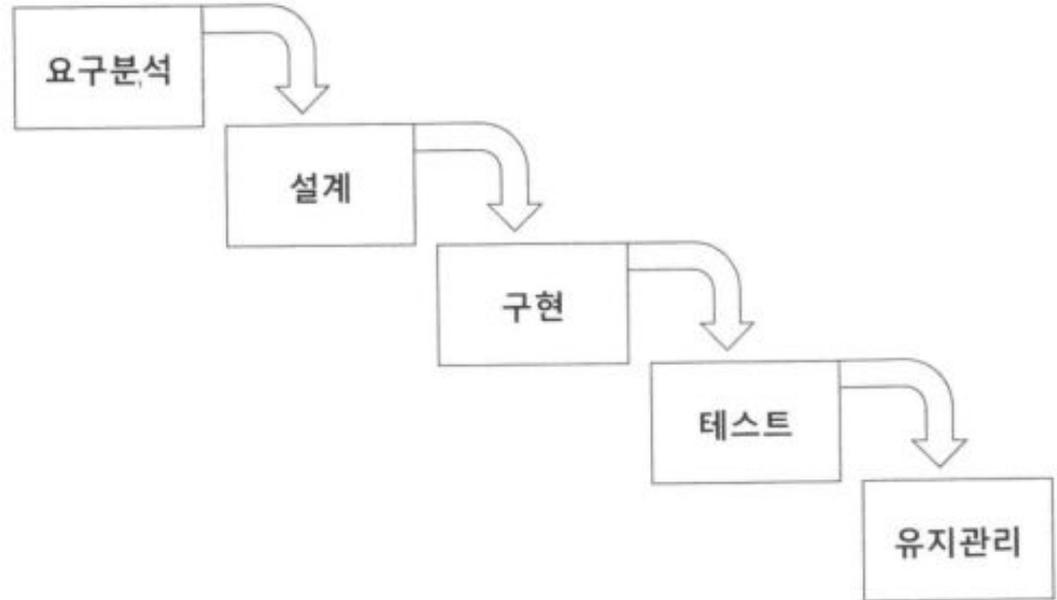


소프트웨어 개발의 위험

- 비가시성(Invisibility)
- 복잡성(Complexity)
- 가변성(Changeability)
- 순응성(Conformity)

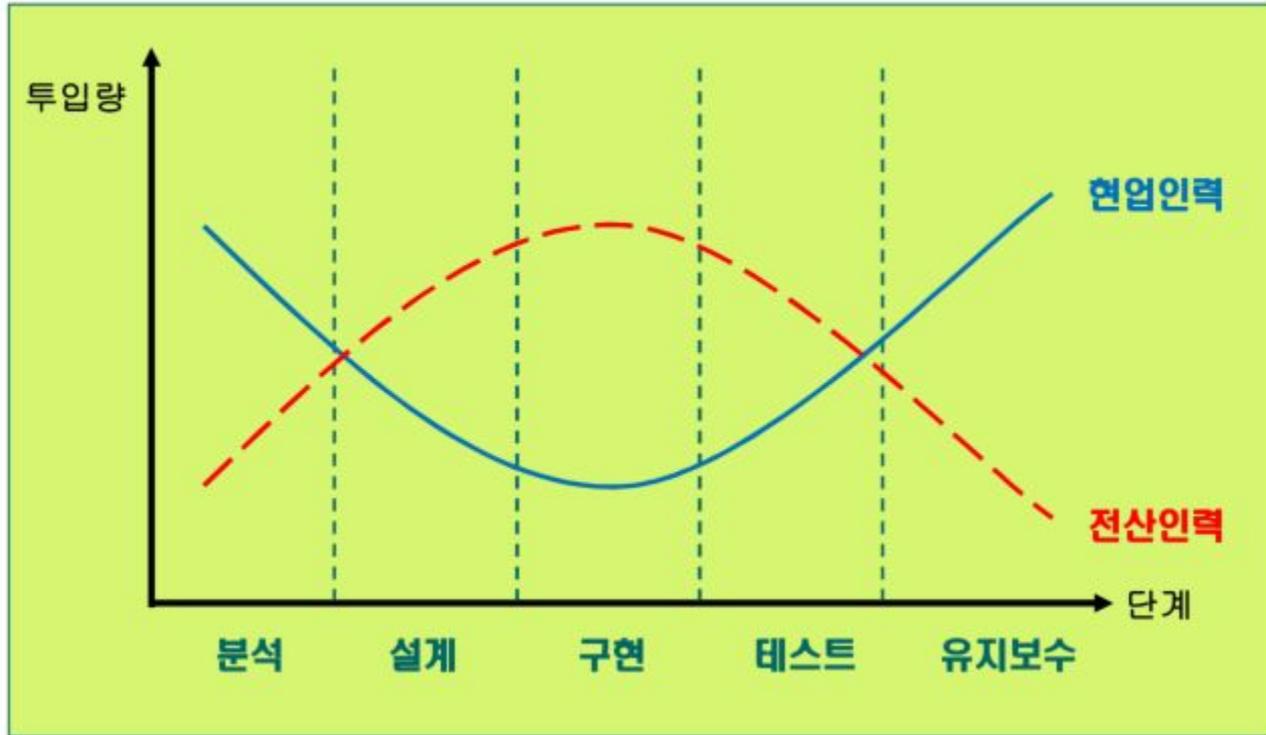
폭포수 모델(waterfall model)

- Requirement Analysis(What)
- Design(How)
- Implementation(Code)
- Test, Debugging
- Maintenance



〈그림 11-2〉 폭포수모델(waterfall model)

개발단계별 인력 소요



신규개발

- 요구사항을 분석하고, 구축방안을 설계하여 새로 개발하는 방법.

조직의 업무상 요구를 시스템으로 개발하여 구현

- 장점 :
 - 조직의 요구를 충실히 반영가능
 - 요구사항 변경 수용 유연성
 - 기존시스템과의 연계 용이
 - 내부 보유인력 활용
 - 예외적인 특수요구 반영
 - 내부 프로세스 변경 최소화
 - 소스코드 소유
- 단점 :
 - 많은 시간과 비용 소요
 - 개발인력의 수준에 품질 좌우

패키지도입(Package)

- 이미 개발된 완제품 소프트웨어를 구입하는 방법으로 가장 손쉬운 방법
조직의 업무를 패키지의 업무절차에 적응시킴.
- 장점 :
 - 개발기간과 비용 절약
 - 즉시 사용 가능
 - 성공사례와 표준화된 절차 수용
 - 자연스러운 업무개선
 - 지속적인 지원활동
- 단점 :
 - 특수한 요구 반영 어려움
 - 기존 절차 변경, 기존시스템과의 연계 어려움
 - 소스코드 확보 어려움

커스터마이징(Customizing)

- 조직의 요구에 맞게 (반제품) 소프트웨어를 변경하여 적용하는 방법

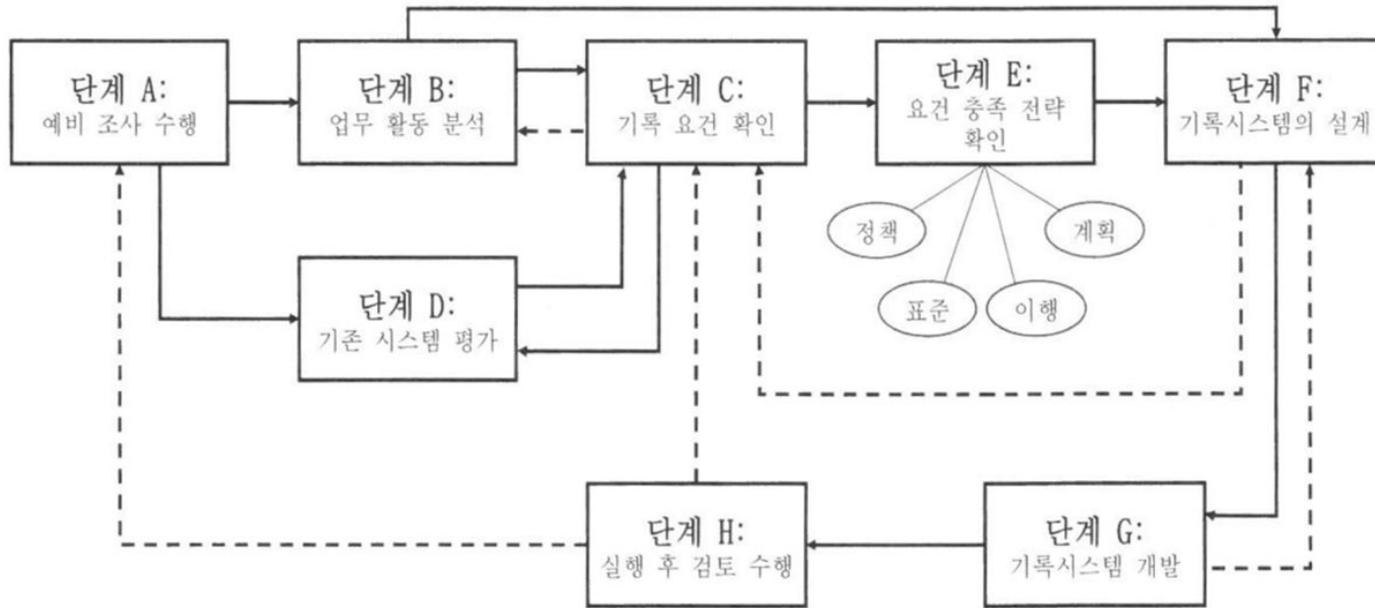
신규개발과 패키지도입의 장점을 결합

- 장점 :
 - 전체 신규개발에 비하여 비용과 기간 절약
 - 신규개발에 비하여 품질 수준 향상 가능
 - 패키지에 없는 특수한 요구사항 반영 가능
- 단점 :
 - 수정된 코드에 대한 변경관리 부담
 - 운영과정에 예기치 못한 문제점이 발생할 수도
 - 무분별한 커스터마이징은 사용자, 공급자 모두에게 위험

사례

- 자료관시스템 (2003~)
- 대통령비서실 기록관리시스템 (2005.1~2007.12)
- 표준기록관리시스템 RMS (2007~)
- 대통령기록관리시스템 PAMS (2007.6~)
- 중앙연구기록관리시스템 CAMS

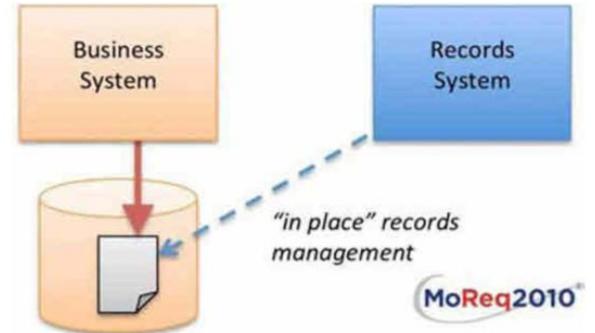
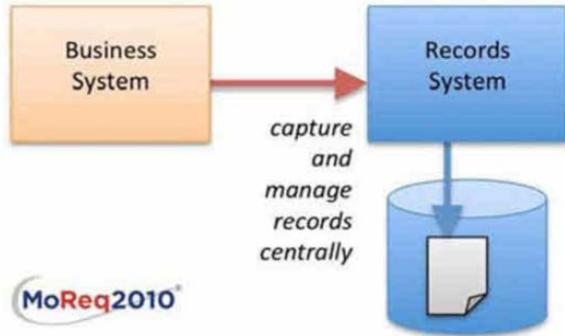
기록시스템의 설계와 구현(DIRS)



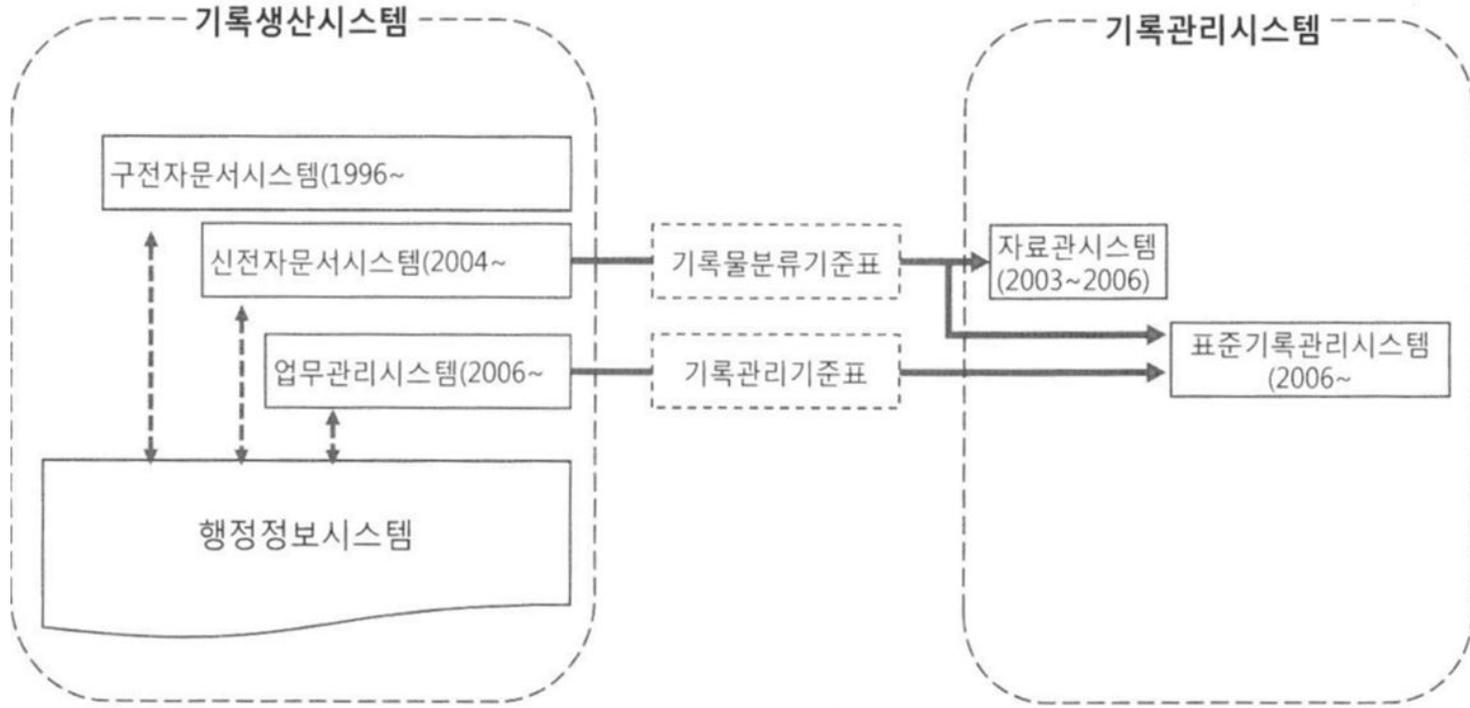
출처: KS X ISO 15489-2 그림 1.

〈그림 11-3〉 기록시스템의 설계와 구현(DIRS)

기록 생산시스템과 관리시스템의 관계



기록 생산시스템과 관리시스템 관계 현황



〈그림 11-8〉 공공부문 기록생산시스템과 기록관리시스템 관계 현황

행정정보시스템

Government-wide Enterprise Architecture

공공부문에서 EA를 공동활용 할 수 있도록 참조모형, 범정부EA, 각 기관의 EA도입 운영 현황 등에 관한 정보를 조회할 수 있습니다.

GEAP 현황 2018년 기준

▪ 범정부 EA 사용기관수	1,257
▪ 정보시스템 수	16,531
▪ 하드웨어 수	213,571
▪ 소프트웨어 수	176,870



법지침



참조모형 / 기준



기타자료실



정보파일목록



장기 보존

장기 보존

- 전자기록 보존의 위험성
 - 비가시성
 - 내용정보와 표현정보의 분리
 - 내용정보와 맥락정보의 분리
 - 빠른 기술 발전속도 (HW, SW 노화)
 - 매체 취약성

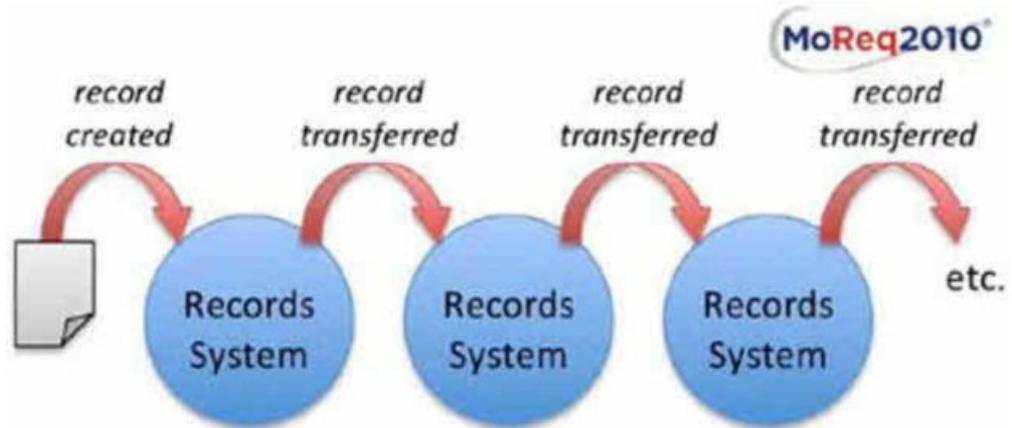
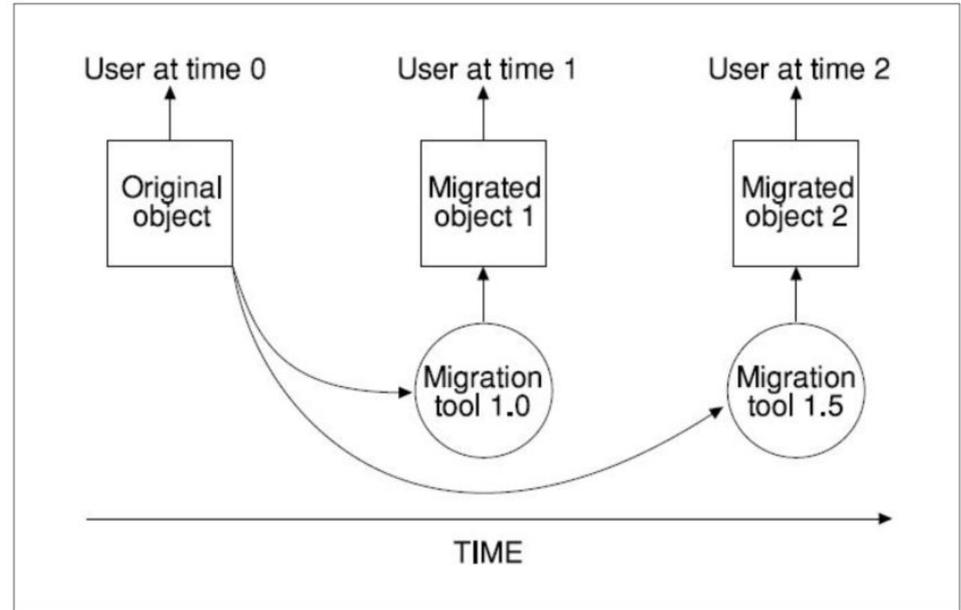


Figure 1g — Records may be transferred multiple times between records systems during their lifespan

장기보존 - 마이그레이션(migration)

- [CCSDS, 2012]
 - 매체 재수록(refreshment)
 - 복제(replication)
 - 재패키징(repackaging)
 - 변환(transformation)
- [ISO 15489-1 : 2016]

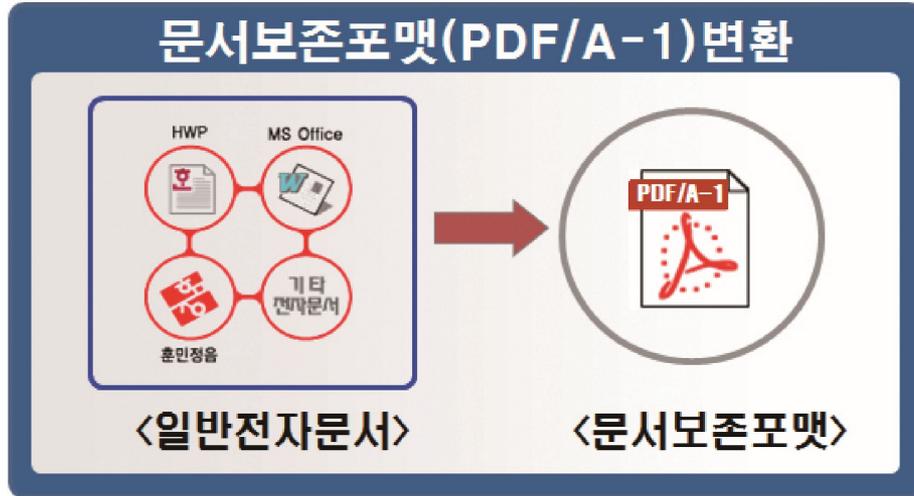
3.13 migration : process of moving records from one hardware or software configuration to another without changing the format



<그림 3-10> Migration on Request의 기본 개념도

장기보존 - 변환(conversion)

3.6 conversion : process of changing records from one format to another [ISO 15489-1 : 2016]



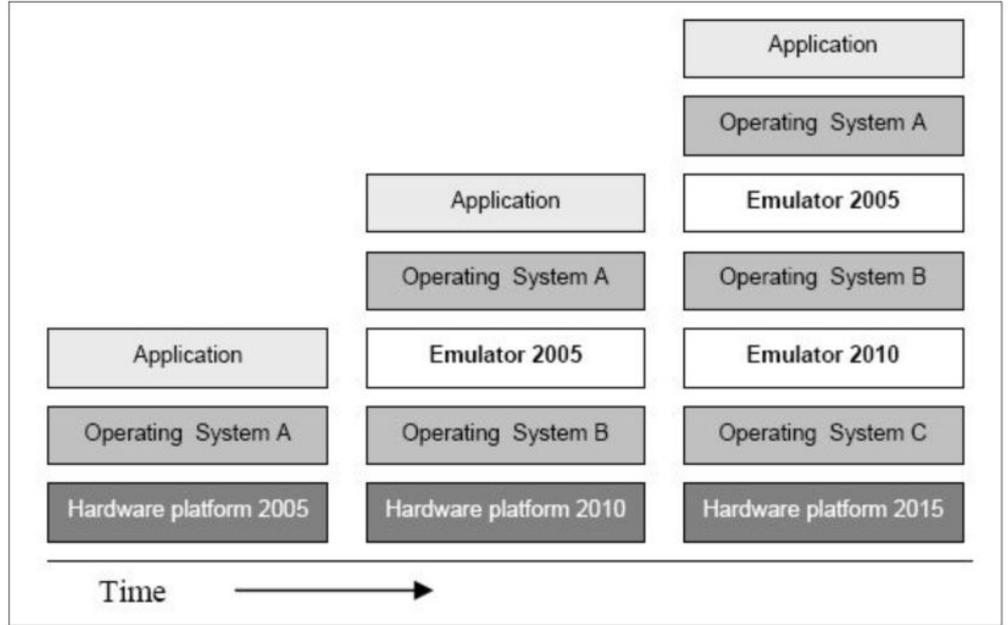
000000001111111100
101101010010101010
1011011000000....



101010100001111000
011111000101101010
010010100001111....

장기보존 - 에뮬레이션(emulation)

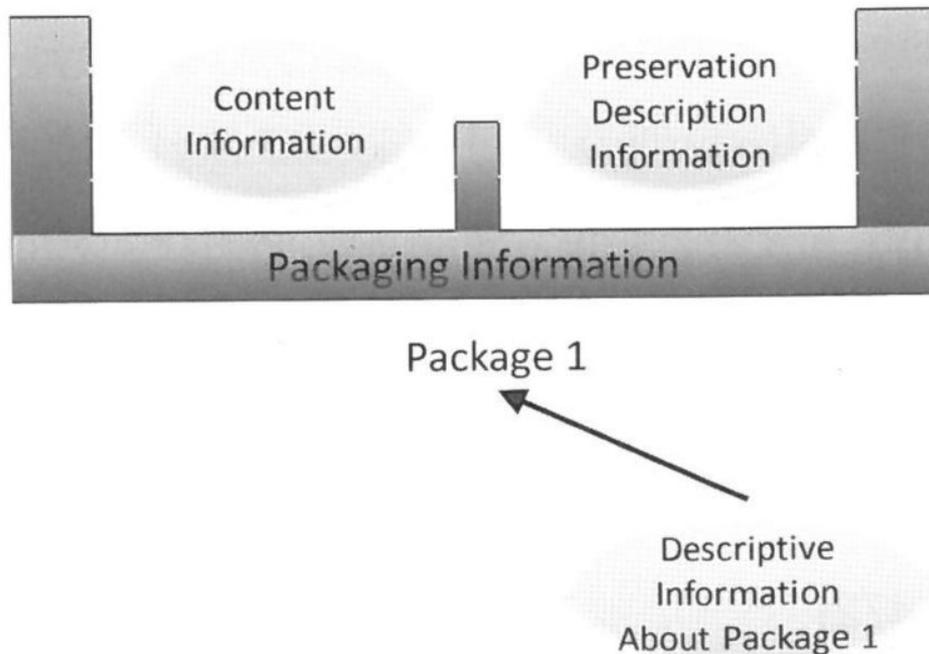
디지털 정보를 생산한 시점에서 사용된 하드웨어, 매체, 운영 체제, 소프트웨어의 운용을 그대로 흉내 내어(emulate) 그 내용을 읽어내는 프로그램을 통하여 재현하는 보존 전략 [기록학용어사전]



<그림 3-8> 누적 에뮬레이션 개념도

장기보존 - 인캡슐레이션(encapsulation)

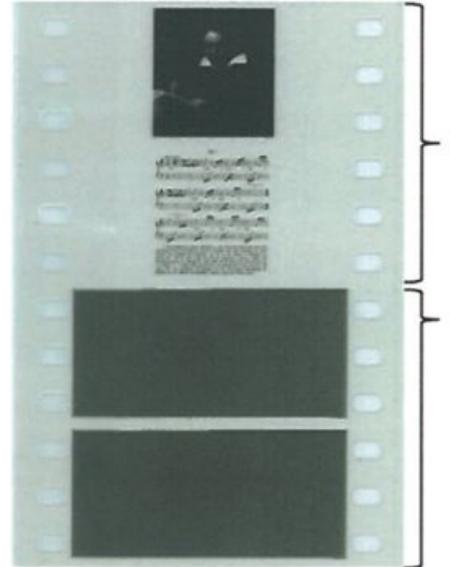
기록을 구성하는 내용과 맥락, 표현정보 등을 하나로 묶어서 보존하는 전략



장기보존 - 아날로그 보존

- 공공기록물관리법시행령 제29조(보존방법) ③항
 - 기록물관리기관의 장은 보존가치가 매우 높은 전자기록물에 대하여는 마이크로필름 등 육안으로 식별이 가능한 보존매체에 수록하여 관리하여야 한다.

- 디지털정보와 육안식별정보를 함께 아날로그 매체에 보존 사례
 - <https://www.piql.com/>



문서보존포맷 - PDF/A-1

- 생산된 당시의 애플리케이션이 없어도 해당문서의 내용과 외형을 그대로 재현하여 내용보기를 가능하게 하는 포맷
- 문서보존포맷 고려요소
 - 공개용 표준
 - 편재성
 - 안정성
 - 메타데이터 지원
 - 상호운용성
 - 진본성
 - 표현력
 - 검색기능
- ISO 19005-1:2005, Document management - Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1)

장기보존포맷 - NEO(개념과 고려사항)

- 오랜 기간이 경과해도 기록물이 생산된 당시에 가지고 있던 내용을 그대로 재현하여 접근할 수 있도록 보존하며, / 업무 활동에 대한 증거와 업무에 대한 책임 소재를 분명하게 밝혀주는 법적 증거를 확보할 수 있도록 / 진본성과 무결성을 보장하면서 장기간 보존
- 장기보존포맷 고려사항
 - 자체 충족성
 - 자체 문서화
 - 구조화된 텍스트 인코딩
 - 무결성

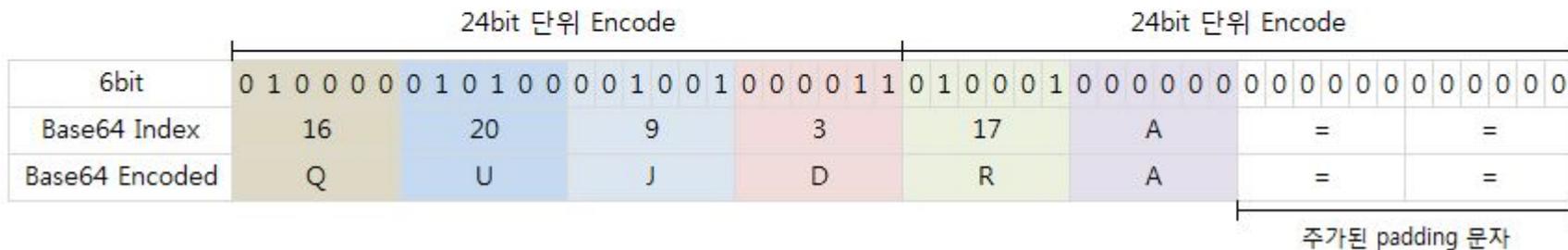
장기보존포맷 - NEO (구성요소와 패키징방법)

- 구성요소
 - 원문
 - 문서보존포맷
 - 메타데이터
 - 전자서명

- 패키징 방법
 - 장기보존포맷의 구성요소를 하나로 패키징할 때 인코딩 방법으로 텍스트 기반인 XML(eXtensible Markup Language)을 이용
 - XML의 특징 : 상호운용성, 가변성, 자체충족성, 자체문서화

장기보존포맷 - NEO (패키징 방법 - base64 Encoding)

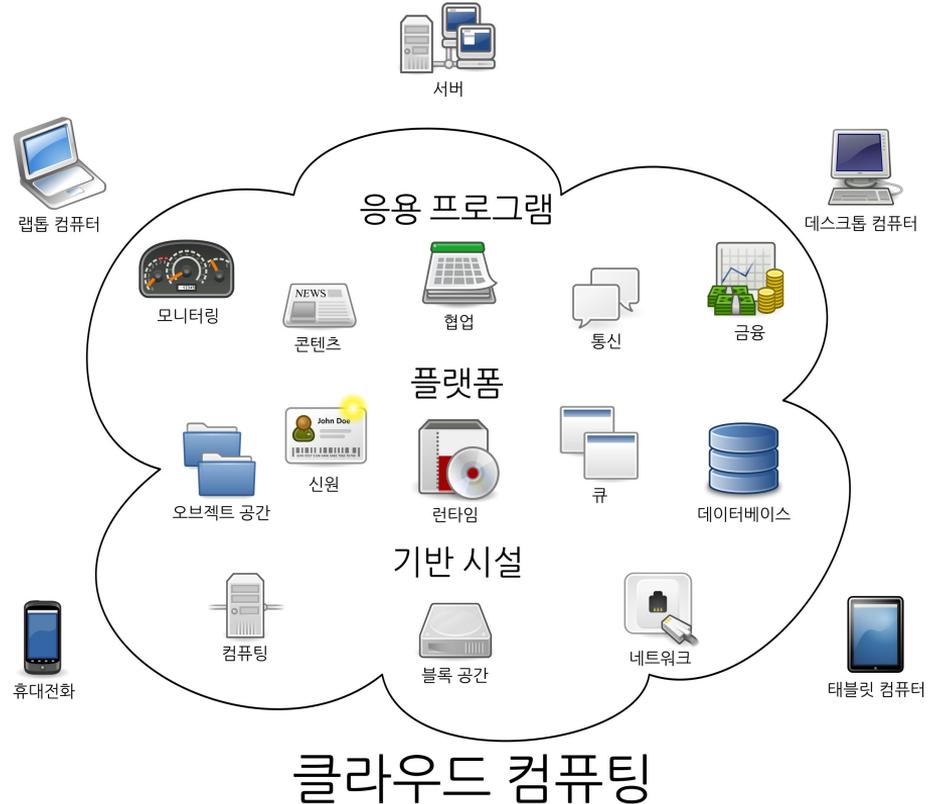
TEXT	A	B	C	D
ASCII	65	66	67	68
8bit	0 1 0 0 0 0 0 1 0 1 0 0 0 0 1 0 0 1 0 0 0 0 1 1 0 1 0 0 0 1 0 0			



그리고...

그리고... 클라우드 컴퓨팅 (cloud computing)

집적 · 공유된 정보통신기기, 정보통신설비, 소프트웨어 등 정보통신자원을 이용자의 요구나 수요 변화에 따라 정보통신망을 통하여 신속적으로 이용할 수 있도록 하는 정보처리체계 *

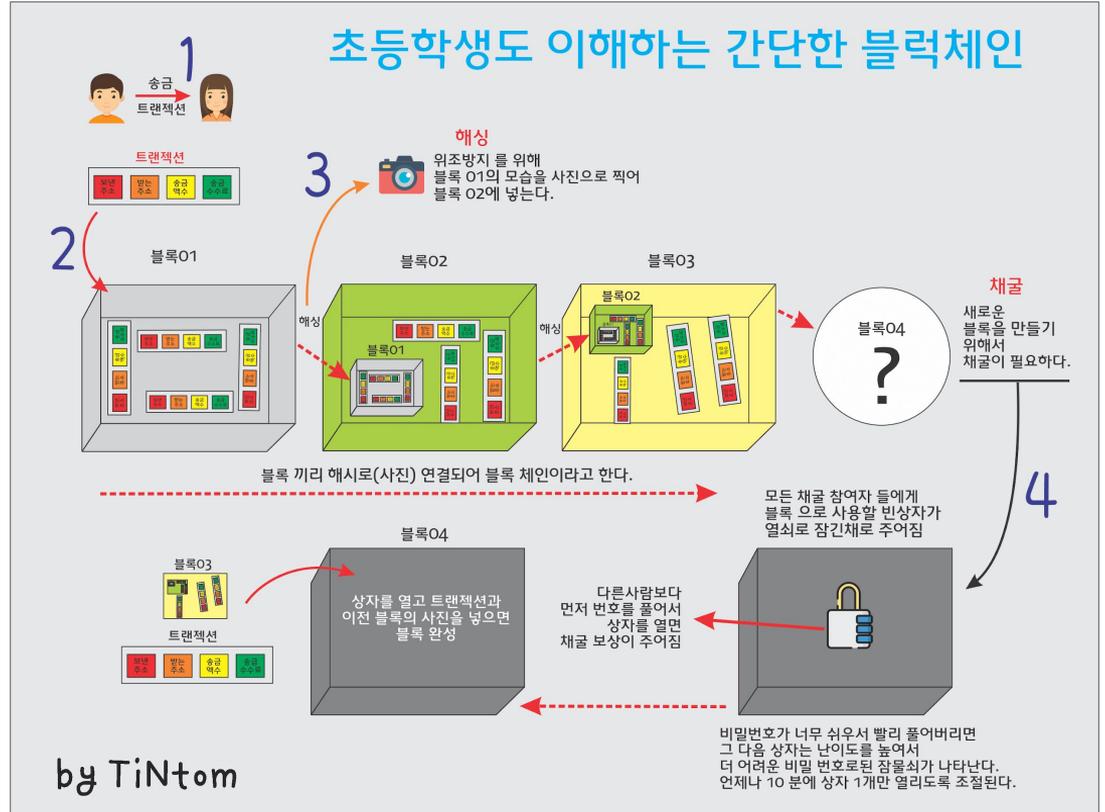


*[클라우드컴퓨팅발전 및 이용자 보호에 관한 법률 제2조(정의)]

[그림출처] https://ko.wikipedia.org/wiki/클라우드_컴퓨팅

그리고... 블록체인 (blockchain)

거래과정의 정보를 참여자들의 네트워크에 분산 저장하여 공동으로 관리함으로써 데이터의 위변조를 방지하는 기술*



*한국기록관리학회(2018), 기록관리의 이론과 실제. 조은글터

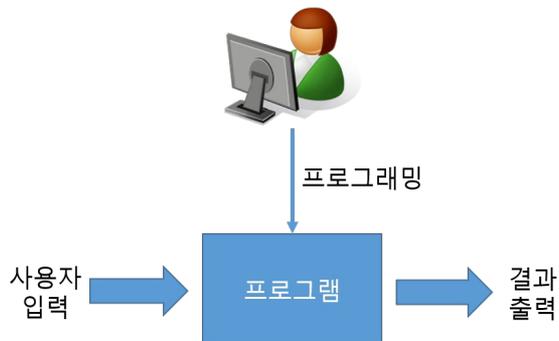
그리고... 마이크로서비스 아키텍처 (microservices architecture)

[그림 2] 마이크로서비스의 기술&조직 구조 독립성

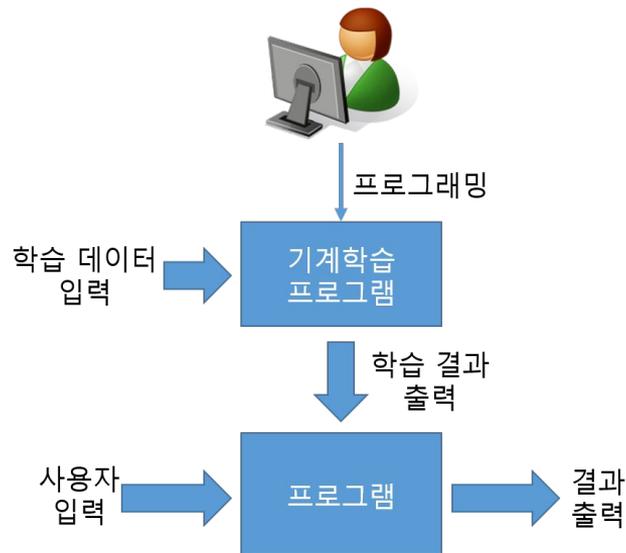


그리고... 기계 학습(machine learning)

-

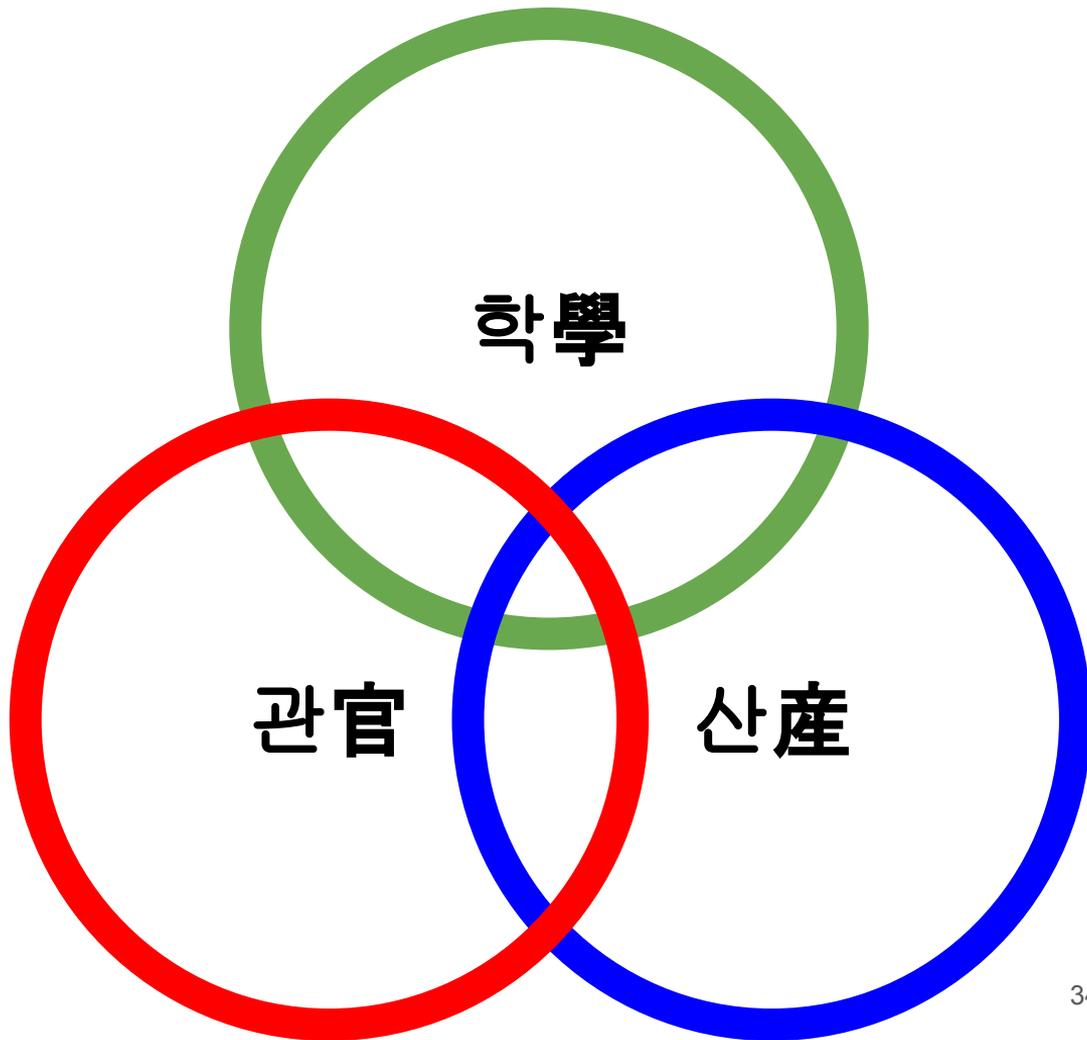


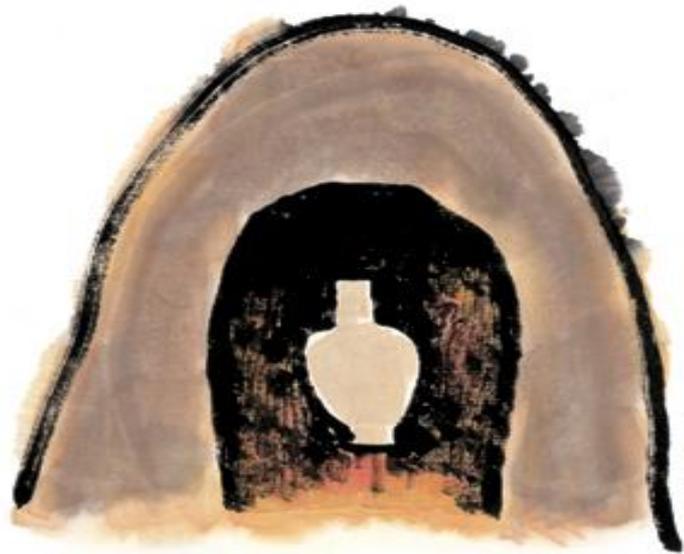
[기존 방식]



[기계학습 방식]

생태계(鼎)





아름다운 도자가
아궁이 앞에 앉아서
생각합니다.
우리와 우리들의
삶을 저마다의
훌륭한
예술품으로
훈도(薰陶)해주는
저마다의 가마를
생각합니다.

이경  