

[별첨 5]

클라우드 기반 데이터맵 상호운용 및 분석 인프라 운영관리 기술



원 회 선 (hswon@etri.re.kr)
스마트데이터연구실



목 차

1. 기술의 개요
2. 기술이전 내용 및 범위
3. 경쟁기술과 비교
4. 기술의 사업성
5. 국내외 시장 동향

1. 기술의 개요

기술 개발 배경

데이터 개방

- 빅데이터에 대한 관심은 증가하나, 활용 가능한 데이터는 부족
- 데이터 산업 활성화를 위해 데이터 개방, 공유, 연계 활성화 필요
- 데이터 마켓/데이터 브로커 등장 등 데이터 유통, 서비스 시장의 태동
- ※ 전체 분석과정 중 데이터 수집, 융합 가공, 준비과정에 최대 80% 시간과 노력 소요 <Gartner, 2015>

데이터 품질

- 원천 데이터 제공, 정확성, 품질관리 미흡으로 실제 활용 미흡
- 데이터의 분류, 형식, 품질 관리 등 종합적인 데이터 관리체계 필요
- 최근 오픈 포맷, 연계, 품질관리 등 데이터 거버넌스 기술 관심 급증
- ※ 빅데이터 성공은 품질이 좌우: 고품질 데이터 보유기업의 연성장률 14%, 후진적 데이터 품질 보유 기업 2% <Aberdeen Group, 2016>

데이터 활용

- 데이터 분석 전문가 부족 → 제한된 공급자 중심의 분석, 지연된 서비스
- 미국, 유럽 등에 비해 국내는 분석 전문가 양성체계 미흡
- 손쉬운 분석 도구, 다양한 API 제공을 통한 사용자 중심 환경으로 전환 필요
- ※ 빅데이터에서 가장 필요한 인력인 데이터 분석 전문가 수요 급증은 주요 이슈 <NIA, 2015>

쓸만한 데이터는 어디에?



믿을 수 없는 데이터,
믿을 수 없는 결과..

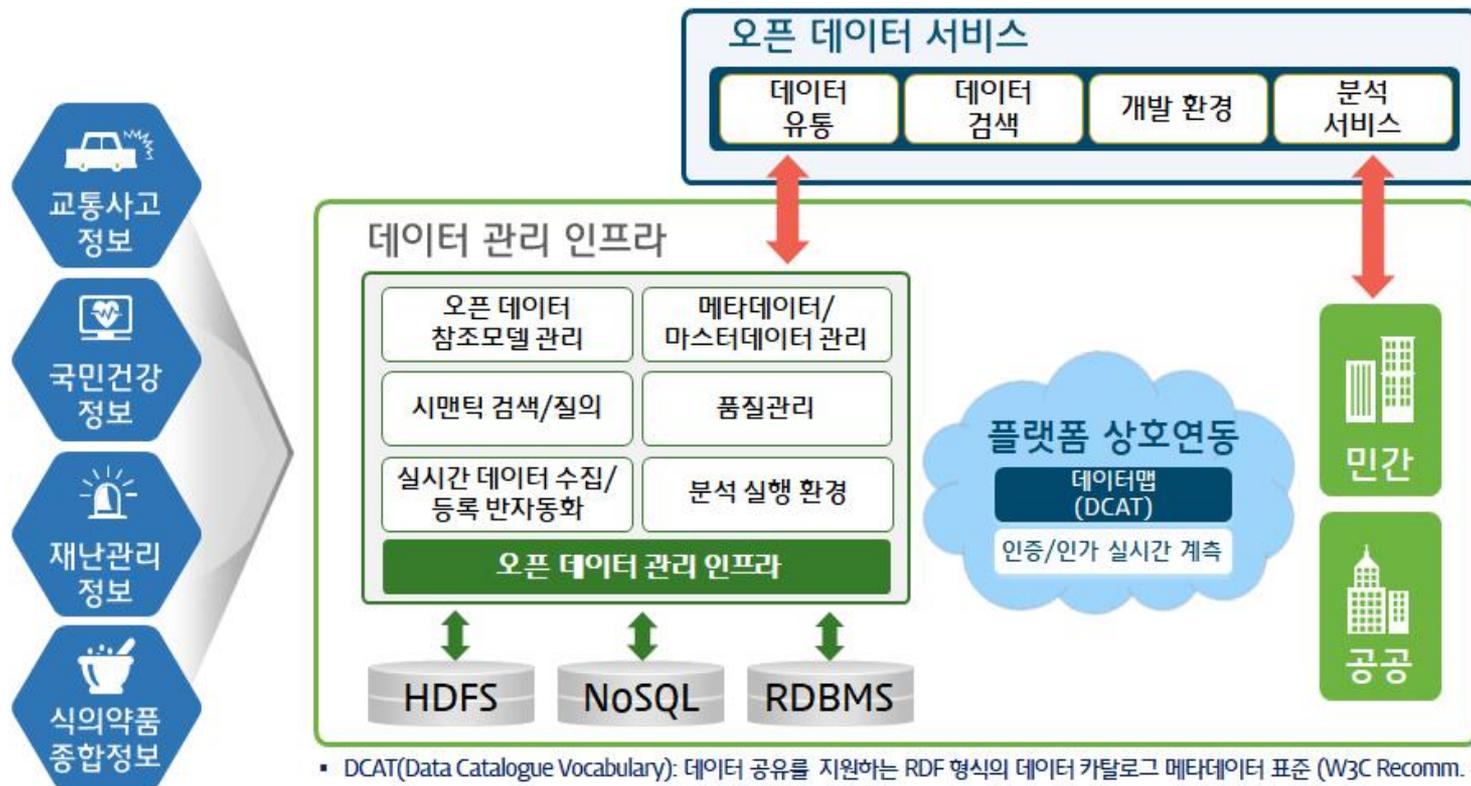


전문가는
부족하고!

1. 기술의 개요

최종 목표

- 누구나 쉽게 데이터 개방하고 공유하기 위한 오픈 데이터 플랫폼 구축에 필요한 핵심 기술 개발



▪ DCAT(Data Catalogue Vocabulary): 데이터 공유를 지원하는 RDF 형식의 데이터 카탈로그 메타데이터 표준 (W3C Recomm. 16)

2. 기술미전 내용 및 범위

□ 데이터맵 상호운용

- ❖ 테넌트(조직 및 사용자) 관리 기술
- ❖ 데이터 분류체계 및 카테고리 관리 기술
- ❖ DCAT 기반 데이터 정보 공유 기술
- ❖ 데이터 유통 기술

□ 멀티테넌트 분석 인프라 운영관리

- ❖ 클라우드 기반 실행환경 서비스 기술
- ❖ SW 이미지(개발환경, 분석도구 등) 프로비저닝 관리 기술
- ❖ 분석 어플리케이션 등록, 배포 및 실행관리 기술
- ❖ 분석 어플리케이션(알고리즘, 오픈 API) 유통 기술

2. 기술미전 내용 및 범위

▣ 기술 개발 현황

❖ 기술성숙도(TRL : Technology Readiness Level) 단계 :

구분	단계	정의	세부설명
기초 연구 단계	1	기초 이론/실험	기초이론 정립 단계
	2	실용 목적의 아이디어/특허 등 개념정립	기술개발 개념 정립 및 아이디어에 대한 특허 출원 단계
실험 단계	3	실험실 규모의 기본성능 검증	실험실 환경에서 실험 또는 전산 시뮬레이션을 통해 기본성능이 검증될 수 있는 단계 개발하려는 부품/시스템의 기본 설계도면을 확보하는 단계
	4	실험실 규모의 소재/부품/시스템 핵심성능 평가	시험생품을 제작하여 핵심성능에 대한 평가가 완료된 단계 3단계에서 도출된 다양한 결과 중에서 최적의 결과를 선택하려는 단계 컴퓨터 모사가 가능한 경우 최적화를 완료하는 단계
시작품 단계	5	확정된 소재/부품/시스템 시작품 제작 및 성능 평가	확정된 소재/부품/시스템의 실험실 시작품 제작 및 성능 평가가 완료된 단계 개발 대상의 생산을 고려하여 설계하나 실제 제작한 시작품 샘플은 1~수개 미만인 단계 경제성을 고려하지 않고 기술의 핵심성능으로만 볼 때, 실제로 판매가 될 수 있는 정도로 목표 성능을 달성한 단계
	6	파일럿 규모 시작품 제작 및 성능 평가	파일럿 규모(복수 개~양산규모의 1/10정도)의 시작품 제작 및 평가가 완료된 단계 파일럿 규모 생산품에 대해 생산량, 생산용량 불량을 등 제시 파일럿 생산을 위한 대규모 투자가 동반되는 단계 생산기업이 수요기업 적용환경에 유사하게 자체 현장테스트를 실시하여 목표 성능을 만족시킨 단계 성능 평가 결과에 대해 가능하면 공인인증 기관의 성적서 확보
실용화 단계	7	신뢰성평가 및 수요기업 평가	실제 환경에서 성능 검증이 이루어지는 단계 부품 및 소재개발의 경우 수요업체에서 직접 파일럿 시작품을 현장 평가(성능 및 신뢰성 평가) 가능하면 인증기관의 신뢰성 평가 결과 제출
	8	시제품 인증 및 표준화	표준화 및 인허가 취득 단계
사업화	9	사업화	본격적인 양산 및 사업화 단계 6-시그마 등 품질관리가 중요한 단계

3. 경쟁기술과 비교

□ 데이터맵 상호운용

- ❖ W3C DCAT 모델 기반 데이터 관리 및 데이터맵 생성, 배포, 수집
- ❖ 기존 경쟁기술 대비 개량된 부분
 - W3C DCAT 확장 버전과 호환되도록 선반영한 데이터 관리체계
 - 미종 플랫폼의 DCAT 기반 데이터맵 생성 지원
 - 데이터 마켓플레이스 구축을 위한 유통 기능

□ 멀티테넌트 분석인프라 운영 관리

- ❖ 클라우드 환경의 분석응용 개발/운영 환경(DevOps) 제공
- ❖ 기존 경쟁기술 대비 개량된 부분
 - 조직, 사용자별 컴퓨팅 자원 할당 제어 및 모니터링
 - 샌드박스 생성 및 SW (개발환경) 프로비저닝 기능
 - 물리노드, 가상머신, 샌드박스 등의 모니터링 기능

4. 기술의 사업성

▣ 예상 응용 제품 및 서비스

예상 제품/서비스	예상 수요자 (층)
데이터 유통 플랫폼	기업, 공공기관
클라우드 분석개발/운영환경 (DevOps)	일반사용자, 기업, 공공기관

▣ 상용화 가능성

- ❖ 각 세부 기능의 모듈화와 API/SDK 제공으로 시장 요구사항에 빠르게 대응하여 커스터마이징 또는 개발 가능
- ❖ 여러 민간 클라우드 포팅이 용이함
- ❖ 조직/사용자의 데이터 관리, 활용과 컴퓨팅 자원 할당 등에 대한 통합 관리 및 모니터링 기능 제공으로 운영 비용 절감

4. 기술의 사업성

□ 기술미전 업체 조건

- ❖ 제품화 소요 예상 시간 (총 12개월)
- ❖ 기술능력: 클라우드 SW 개발, 데이터 모델링 및 DB 기술 개발자 보유
- ❖ 소프트웨어 구현 및 디버깅을 위한 개발환경은 실시권자가 구축

□ 사업화시 제약 조건

제약 조건	극복(개선)방안
국내외 표준체계 정립단계	표준개발 현황 추적 및 데이터 모델링
실증 테스트 환경 구축	데이터 보유 및 포털 운영기관 협업 추진

5. 국내외 시장 동향

□ 시장 동향

- ❖ 세계 데이터 시장규모는 ' 17년 1,508달러에서 ' 20년 2,100억 달러로 연 11.9% 성장 전망

(출처: Worldwide Semiannual Big Data and Analytics Spending Guide, IDC 2017.04)

- ❖ 국내 데이터 시장규모는 ' 17년 6조 2,973억원에서 ' 20년 7조 8,450억원으로 ' 22년에는 10조원으로 연7.6%씩 성장미 예상됨

(출처: 2017년 빅데이터 실태 조사)

□ 시장성

예상 제품 /서비스	예상단가 (천원)	이전기술의 비중(%)	잠재적/현재적 경쟁자와 가격, 시장 등에서 경쟁상 유리한 점	판매가능 시기
데이터 유통 플랫폼	200,000	50	a. 가격경쟁력면: (기존제품 대비 가격 50%절감) b. 시장환경면: (예시: <u>2021년</u> 이후 연간 1천억 예상)	2021
클라우드 개발/배포 환경(DevOps)	200,000	50	a. 가격경쟁력면: (기존제품 대비 가격 50%절감) b. 시장환경면: (예시: <u>2021년</u> 이후 연간 1천억 예상)	2021

감사합니다.



www.etri.re.kr