



오픈소스 빅데이터 기술 적용사례

2017.07.12

ONYCOM

어니컴 주식회사 이성준 이사

1. 구축사례 개요

※ 본 내용은 공고된 제안요청서 기반 재구성한 내용 임

| | |
|---------|--|
| 사업명 | 빅데이터 활용 식품 사고·위해 예측기반 구축 |
| 수요기관 | 식품의약품안전처 정보통계담당관실 |
| 사업기간 | 2016.08 ~ 2016.12 (4.5개월) |
| 사업예산 | 680백만원 이내 (VAT 포함) |
| 사업목적/목표 | <ul style="list-style-type: none">▪ (사업목적) 빅데이터 활용 사고·위해 예측 기반 구축으로 식품위기 선제대응 강화 및 식품안전관리·정책 의사결정의 정확성 제고▪ (사업목표) 식품 사고·위해 사전차단을 위한 빅데이터 예측 기반 구축 |
| 사업내용 | <ol style="list-style-type: none">1. 식품 사고·위해 예측모델 구축2. 빅데이터 분석 체계 응용 프로그램 개발3. 빅데이터 플랫폼 설계 및 구축 |

2. 수행내용 및 절차

※ 본 내용은 공고된 제안요청서 기반 재구성한 내용 임



빅데이터 분석플랫폼 구축

데이터 요구사항
수집 및 분석

데이터 연계 및
통합 DB 구축

빅데이터 분석시스템 구축

식품 빅데이터 활용 포털 서비스 구축

식품 위해 예측서비스

요구사항 분석 및 데이터 수집

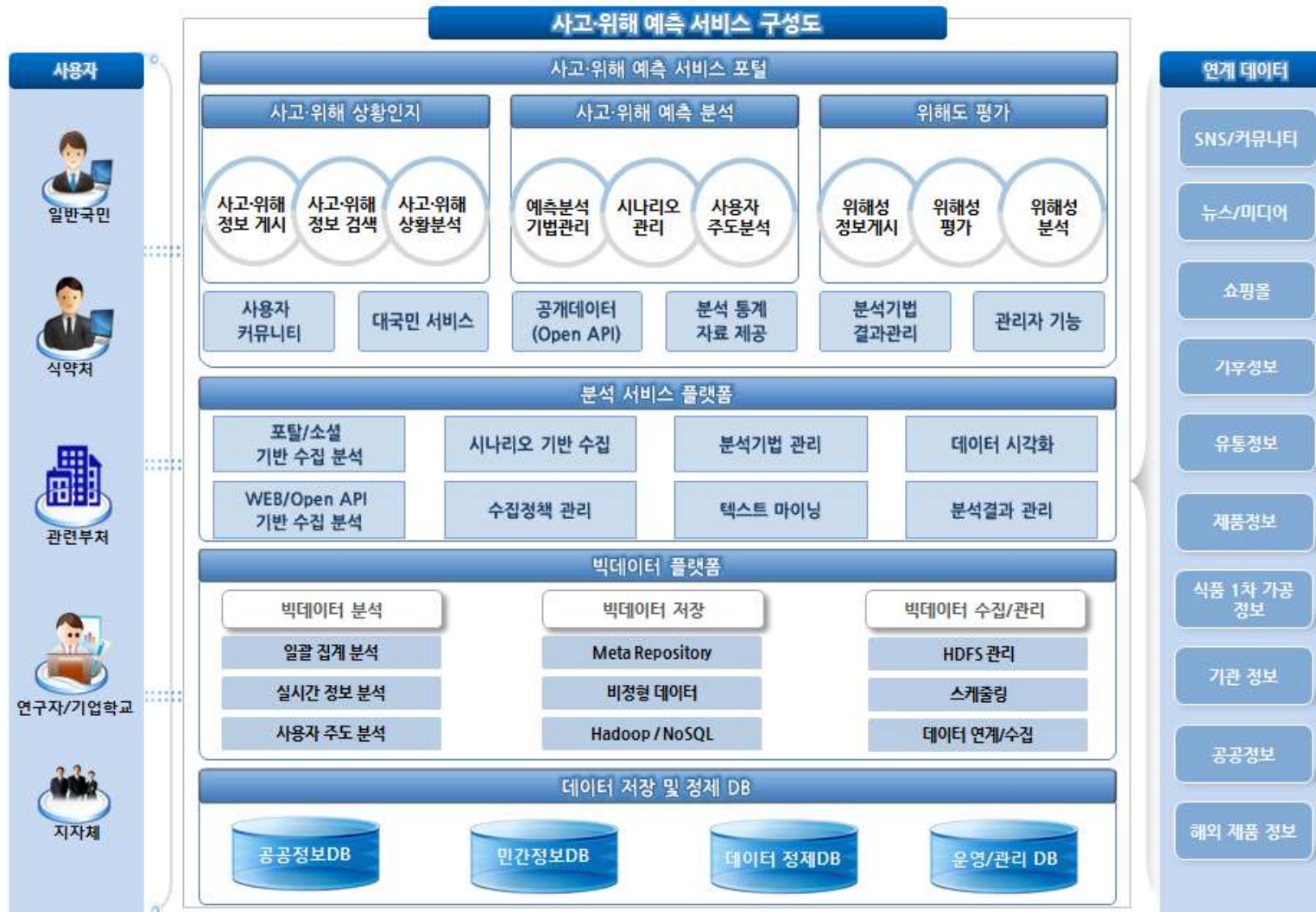
통계/데이터마이닝 기반 예측모델 생성

예측결과 서비스 적용

3. 수행내용 상세

3-1. 서비스 구성도

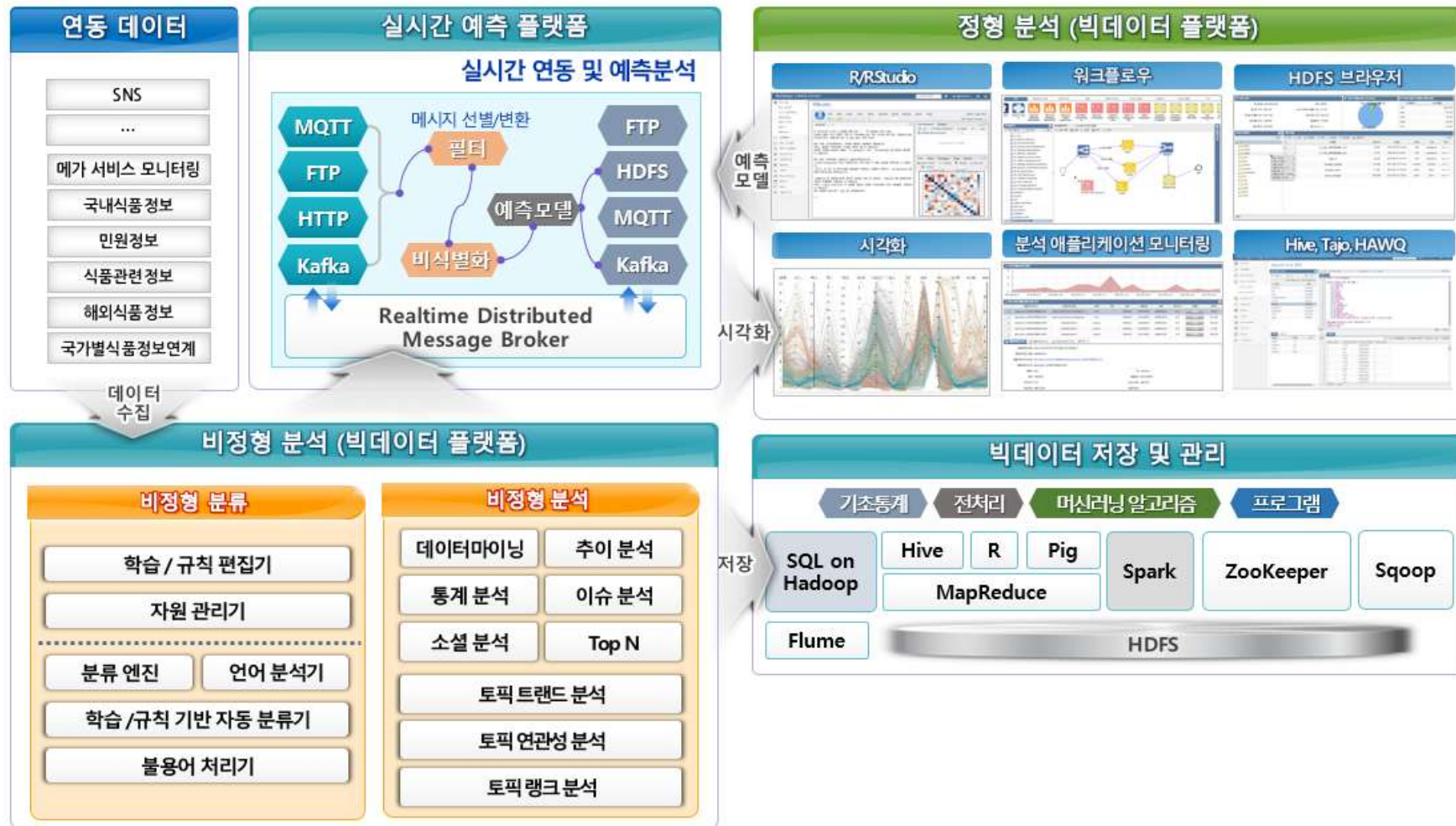
※ 본 내용은 공고된 제안요청서 기반 재구성한 내용 임



3. 수행내용 상세

3-2. 빅데이터 분석 플랫폼 구성도

※ 본 내용은 공고된 제안요청서 기반 재구성한 내용 임



4-1. ankus Analyzer : 빅데이터 분석 플랫폼

※ 본 내용은 ankus analyzer 최신 본으로 납품 본과 일부 상이함

Hadoop 기반 빅데이터 분석 플랫폼

(웹 브라우저 지원, 손쉬운 드래그 & 드롭 사용 지원, 타 OSS 탑재 · 통합 지원)

The screenshot displays the ankus Analyzer web interface. At the top, there is a navigation menu with tabs: 전체, 통계분석, 전처리, 연관규칙분석, 분류, 군집, 기타. Below the menu is a toolbar containing various analysis tool icons with labels: Numeric Statistics (Distribution), Nominal Statistics (Distribution), Certainty Factor based Sum, TFIDF, DOCSIMILITY, Min/Max Normalization, ETL, EM Discretization, MDL Discretization, FP-Growth, ID3 Classification, C4.5 Classification, Naive Bayes Classification, MLP, KNN Classification, GraphAnalysis, K-Means Clustering, and EM Clustering. The main workspace shows a workflow editor with a '하둡 클러스터' dropdown set to 'noflamingo'. Below this are buttons for 상세설명, 신규, 저장, 실행, and XML보기. The workflow diagram consists of a 'Start' node, a 'C4.5' decision tree node, and an 'End' node, all connected by arrows. On the left side, there is a '워크플로우' section with a list of workflow items: c45, numeric, hamoni01, and connect.

4. 활용 OSS



4-2. ankus : 빅데이터 분석 토탈 솔루션

1) 빅데이터 분석 토탈 솔루션 : ankus

ankus logo | nipa 정보통신산업진흥원 National IT Industry Promotion Agency 공개SW 개발 R&D 과제 선정 (2012.04~2013.11) | NIA 한국정보화진흥원 빅데이터 우수솔루션 상품화 지원과제 선정 (2014.12) | 중소기업청 Small & Medium Business Administration 고도화 개발 R&D 과제 선정 (2014.11~) | Go to Global!

2) ankus 제품 구성 : 오픈소스(Community Ver.)와 상용버전(Commercial Ver.) 구성

| | |
|--|--|
| <p>www.openankus.org (OSS 버전, 커뮤니티 주관)</p> | <p>www.openankus.com (상용 버전, 어니컴 주관)</p> |
|  <p>ankus logo</p> |  <p>ankus logo</p> |

4-2. ankus : 빅데이터 분석 토털 솔루션



Hadoop 기반 BigData Total Solution 브랜드

ankus(앵커스)는 코끼리 조련용 도구를 의미하는 인도어로, '빅데이터를 분석하는 도구'의 의미를 담고 있는 빅데이터 토털 솔루션 브랜드입니다.
 <분석기술(Library), 분석제품(Analyzer/Optimizer/Appliance), 분석서비스(4 Industry)로 구성>

| 제품/서비스명 | 배포처 | 라이선스 | 기능 | 제품 지원 | |
|------------|-------------------------|--------|--|-------|----|
| | | | | 공개 | 상용 |
| Library | github.com/openankus | GPL v3 | 데이터 마이닝/머신러닝 라이브러리 (공개 10종, 상용 26종) | ○ | ○ |
| Analyzer | github.com/openankus | GPL v3 | 데이터 마이닝/머신러닝 분석 플랫폼 (공개 20종, 상용 26종) | ○ | ○ |
| Optimizer | github.com/openankus | GPL v3 | 예측모델 최적화 모듈 | ○ | - |
| Crawler | Github.com/onycom-ankus | Apache | 웹 크롤링 (Http Components Hclient 4.5 활용 개발) | ○ | - |
| Appliance | - | - | 빅데이터 어플라이언스 (HW+SW+SVC.) | - | ○ |
| 4 Industry | - | - | 산업별 지능화 서비스 구현 (데이터+알고리즘+도메인) | - | ○ |

5. 활용시 고려사항

5-1. 오픈소스 관점

1 경쟁력 있는 국내産 OSS의 활용 추천

➢ 외산 OSS와 어깨를 겨룰 수 있는 우수한 국내 개발 OSS 활용 검토

● 국내 개발 OSS



● 국내외 OSS 결합 활용 (ankus+R)



5-1. 오픈소스 관점

2

지속적인 기술지원을 위한 '기술지원 예산' 고려

➤ OSS는 초기 셋업 못지 않게 지속적인 업데이트/기술지원 필요

● OSS의 장단점

| 장점 | 단점 |
|---|---|
| <ul style="list-style-type: none">▪ 무료 OSS를 활용한 라이선스 비용절감▪ 특정 SW 및 HW의 종속성 탈피▪ 기존에 개발된 OSS 소스코드를 이용한 빠른 개발▪ 오픈포맷/표준화 프로토콜 활용에 따른 유연한 연동성▪ 전세계 개발자 참여로 인한 빠른 발전속도 및 높은 신뢰성 | <ul style="list-style-type: none">▪ 자유로운 개발에 따른 체계적이지 못한 문서▪ 무료인 OSS 사용에 따른 제품보증 및 유지보수의 어려움▪ OSS 라이선스 미 준수에 따른 저작권/특허 이슈 위험성 |

<출처: 한국저작권위원회>

5-2. 빅데이터 관점

3

빅데이터 플랫폼 도입 시 고려 요인 : 분석 알고리즘

➢ 최근 예측 모델 구축 강제, 지능형 서비스 구현을 위한 알고리즘 중요성 인지 필요

데이터

- 정형데이터
- 비정형 데이터
- 센서 데이터
- 아날로그 데이터
- etc.



알고리즘

- 데이터마이닝
- 머신러닝
- 딥러닝
- 인공지능망
- etc.



지능형서비스



(예측모델) 학습

6. OSS 공개 계획

6-1. 2017년 빅데이터 관련 OSS 공개 계획



6. OSS 공개 계획



6-2. ankus Crawler

The screenshot shows the GitHub repository page for 'onycom-ankus / ankus_crawler'. The repository is currently empty, with 0 stars, 0 forks, and 0 watches. It has 2 commits, 1 branch, and 0 releases. The license is Apache-2.0. The repository was last committed 3 days ago by onycom-ankus. The file list includes:

| File Name | Commit Message | Time |
|---------------------------|----------------------|------------|
| bin/org/ankus | Add files via upload | 3 days ago |
| lib | Add files via upload | 3 days ago |
| src/org/ankus | Add files via upload | 3 days ago |
| LICENSE | Initial commit | 3 days ago |
| Plan of ankus Crawler.pdf | Add files via upload | 3 days ago |
| script.prj | Add files via upload | 3 days ago |
| 수집기 사용 매뉴얼.pdf | Add files via upload | 3 days ago |

The screenshot shows a news article from CIOBIZ titled '어니컴, 웹 기반 문서 수집기 '앵커스 크롤러' 공개' (Ankcom, Web-based Document Collector 'Ankus Crawler' Released). The article is dated 2017.07.10. The main image shows a person interacting with a futuristic digital interface. The article text includes:

어니컴(대표 지영만)은 '앵커스 크롤러(ankus Crawler)'를 아파치 라이선스로 공개했다고 10일 밝혔다.

앵커스 크롤러는 아파치 Http 컴포넌트 프로젝트 'Http 클라이언트 v4.5'를 기반으로 개발한 웹 기반 문서 수집기다. 웹 크롤러에 대한 기본 지식 없이도 간단한 정의수준 자바 프로그래밍과 HTML 해독능력만 있으면 웹 문서 수집이 가능하다.

장점은 스크립트 과일을 이용해 HTML 문서에서 정보추출과 수집대상 HTTP 웹 주소정의가 가능한 점이다. 그래픽이용자환경(GUI) 적용 시 손쉽게 활용 가능하다.

앵커스 크롤러는 오픈소스 전문사이트 깃허브에서 내려받기 가능하다.

[전자신문 CIOBIZ] 김지선기자 river@etnews.com

주소: 서울특별시 영등포구 여의나루로 12가길 51 전자신문 대표이사: 구원모 사업자번호: 107-81-80959 대표전화: 02-2168-9200

Target? 공공기관

When? 2017.08.01

Where? Github

Target? 생산제조기업

When? 2017.11.30

Where? Github



BigData Analysis Total Solution

별첨

- [별첨1] 회사 소개
- [별첨2] 솔루션 소개 : ankus
- [별첨3] 빅데이터/IoT 사업 실적

[별첨1] 회사 소개



1-1. 회사 개요

| ONYCOM | | | |
|--------|--|-----|---------------------|
| 회사명 | 어니컴 주식회사 | 대표자 | 지영만 |
| 설립일 | 1998년 09월 10일 | 업태 | SW자문개발 및 공급 외 |
| 매출액 | 143억 원 (2016년 기준) | 직원수 | 240명 (2017년 06월 기준) |
| 소재지 | [본 사] 서울시 중구 세종대로 21길 22, 태성빌딩 4층 [연구소] 경기도 용인시 기흥구 영덕동 1019 흥덕 U타워 2809호 | | |
| 주요 약력 | <ul style="list-style-type: none"> ▪ Mobile 3G/4G Platform SDK 개발, 단말기 SW개발 Tool 개발 (2002~2008) ▪ 삼성 bada SDK 개발, bada App. 개발 파트너사 선정 (2009~2013) ▪ 솔루션&서비스 개발 및 유통 - Big Data, IoT, Cloud, Security, etc. (2012~) ▪ 협·단체 활동 - (사)스마트개발자협회 회장사(2012~), 빅데이터전문가협의회 부의장사(2013~) ▪ 해외 지사 보유 - ONYCOM USA, Inc.(2006), 베트남 하노이 법인(2012), 일본 도쿄 법인(2015) | | |
| 수상 이력 | '제 11회 모바일 기술대상' 방송통신위원회 위원장상 수상 (2011.11.29) – TestForte 'SK 주식회사 C&C, Biz. Idea Festival 최우수상 수상 (2015.08.05) – 모바일 테스트 센터 | | |
| 주요 인증 |      | | |

1-2. 빅데이터 사업역량

SOLUTION

자체 빅데이터 솔루션 보유

- 기술(1), 제품(3), 서비스(1)로 구성
- 국책 과제 통한 제품 개발/공개 (2013)
- NIA 우수 솔루션 선정 (2014)
- GS 인증 획득 (2016)
- 대한민국 IT브랜드 대상 수상 (2016)

HR

빅데이터 전문인력 보유

- 인공지능/머신러닝 석박사 인력 구성
- 알고리즘&플랫폼 개발/분석가로 구성
- 빅데이터/기계학습/AI/IoT 분야 수행
- 분석기술/제품/서비스 구현 역량 보유

REFERENCE

다양한 수행사례 보유

- 분석과제 발굴 (서울특별시)
- 정보화전략계획 수립 (ISP, 서울특별시)
- 플랫폼 구축 (식약처 외 다수)
 - 식약처(식품), 폴리텍대학(S/Factory), 누리텔레콤(전력), 소프트센(IoT)
- 분석서비스 (서울특별시 외 다수)
 - 서울특별시, 산림과학원, 식약처, 누리텔레콤

튼실한 기업 업력

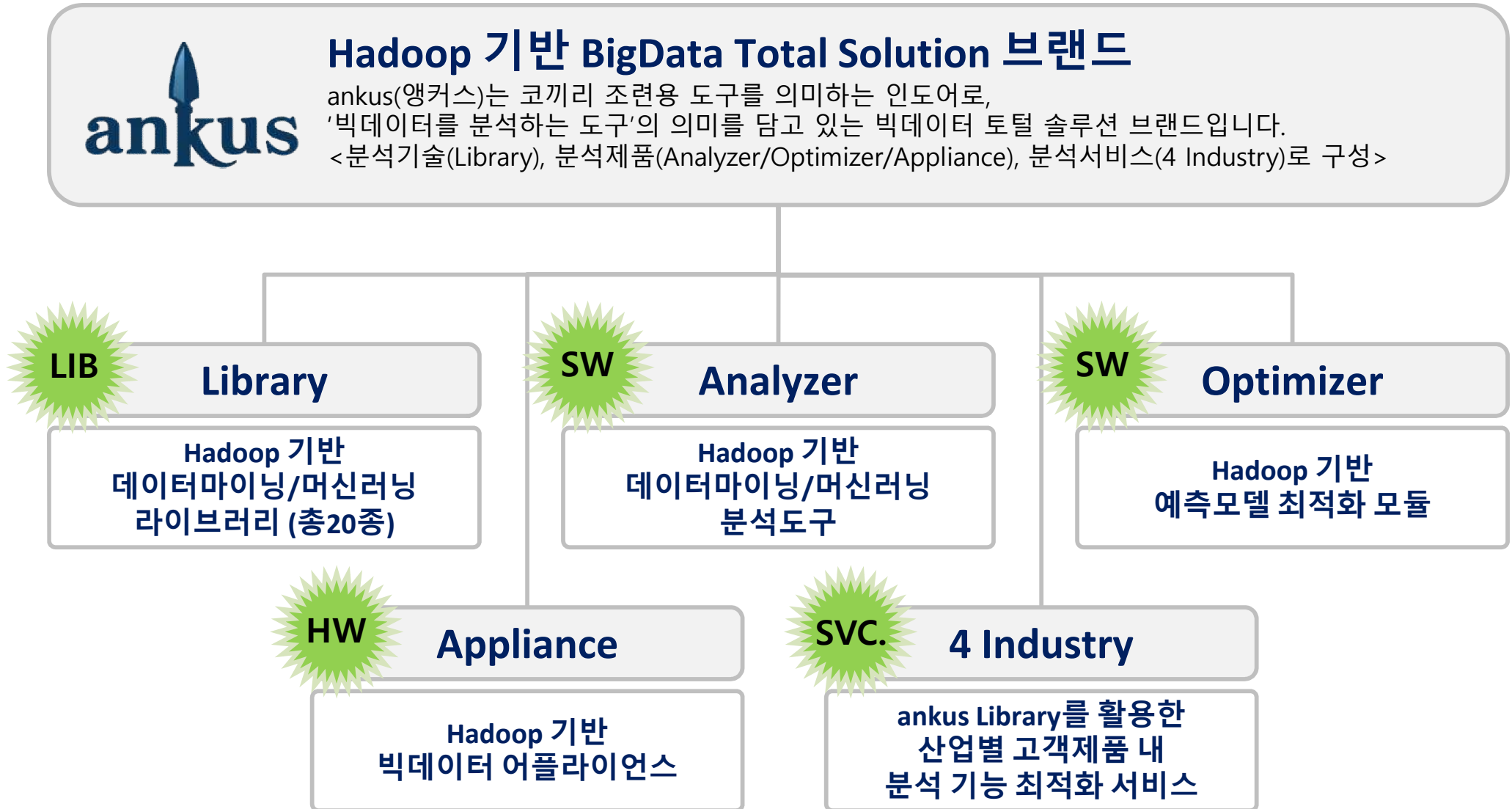
- 업력 19년 차 ICT 중소기업

COMPANY

안정적인 기업 규모

- 연 매출 143억 원, 임직원수 240명 규모

2-1. 제품 체계



2-2. ankus History

2013

- [04월] ankus 개발 계획, 공개SW개발지원사업 선정
- [12월] ankus 0.1 Source Code Open

2014

- [02월] ankus, 서울시 과제(택시매치메이킹) PoC 분석 활용
- [11월] ankus 고도화 개발 계획, 혁신기업기술개발사업 선정
- [12월] ankus, 미래부 빅데이터 우수솔루션 상품화 과제 선정

2015

- [01월] ankus, NIA K-ICT 빅데이터 센터 분석 알고리즘 적용
- [08월] ankus, SW수출 마케팅 지원사업 선정, 일본 진출
- [10월] ankus 4 IoT 서비스 출시
- [10월] ankus 1.0 Release

2016

- [02월] ankus, Custom License 상품 출시
- [03월] ankus, 삼양데이터시스템과 전략적 MOU 체결
- [06월] ankus Optimizer, 공개SW개발지원사업 선정
- [09월] ankus analyzer, GS 인증 획득
- [11월] ankus, 대한민국 IT 브랜드 대상 수상
- [11월] ankus, IT JAPAN WEEK 秋 전시회 출전
- [12월] ankus Optimizer Source Code Open

2017

- [06월] 전통산업 공개SW요소기술지원사업 선정
- [07월] ankus Crawler Source Code Open

| | |
|-------------------------------------|--|
| <p>1</p> <p>ankus Library</p> | <p>K-ICT 빅데이터센터</p> <p>분석 인프라 내 Library 적용 (2015.01)</p> <hr/>  <p>금융상품 추천 지능형 온라인 자산관리 시스템 (2016.09~10)</p> |
| <p>3</p> <p>분석 컨설팅</p> | <p> 서울시 빅데이터 공유활용 기반 조성사업 ISP 컨설팅 (2013.10~2014.02)</p> <p> 소각산불 빅데이터 수집 및 분석 (국립산림과학원, 2014.10~2014.12)</p> <p> 빅데이터 활용 식품사고 위해 예측 모델 구축 (식약처, 2016.09~12)</p> <p> 에너지 사용량 예측 (누리텔레콤, 2017.01~02)</p> |
| <p>2</p> <p>ankus analyzer</p> | <p> (주)소프트센 SOFTCEN CO., LTD.</p> <p>IoT 플랫폼 (2016.01~02)</p> <hr/> <p> 식품의약품안전처</p> <p>빅데이터 분석 플랫폼 (2016.09~12)</p> <hr/> <p> NURI TELECOM</p> <p>에너지 IoT 플랫폼 (2017.01~02)</p> |
| <p>4</p> <p>ankus appliance</p> | <p>※ OSS ankus 사용 기관/대학</p> <ul style="list-style-type: none"> - [기관] 도로교통공단, 생산기술연구원, 육군본부 外 - [대학] 세종대학교, 강원대학교 外 <p> 한국폴리텍대학 울산캠퍼스</p> <p>한국폴리텍대학 울산캠퍼스 (2016.10)</p> |

감사합니다.

ONYCOM

어니컴 주식회사 (ONYCOM, Inc.) 이성준 이사

Address. 서울특별시 중구 세종대로21길 22, 태성빌딩 4층

Tel. 070-4620-6011 / 010-4228-8362 / Fax. 02-541-0090

E-mail. leesj@onycom.com

Homepage. www.onycom.com, www.openankus.com

