

[별첨 5]

보급형 KSB 인공지능 프레임워크 기술



이연희 (leeyh@etri.re.kr)
자가학습엔진연구실

목 차

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

1. 기술의 개요

□ 이전기술의 개요

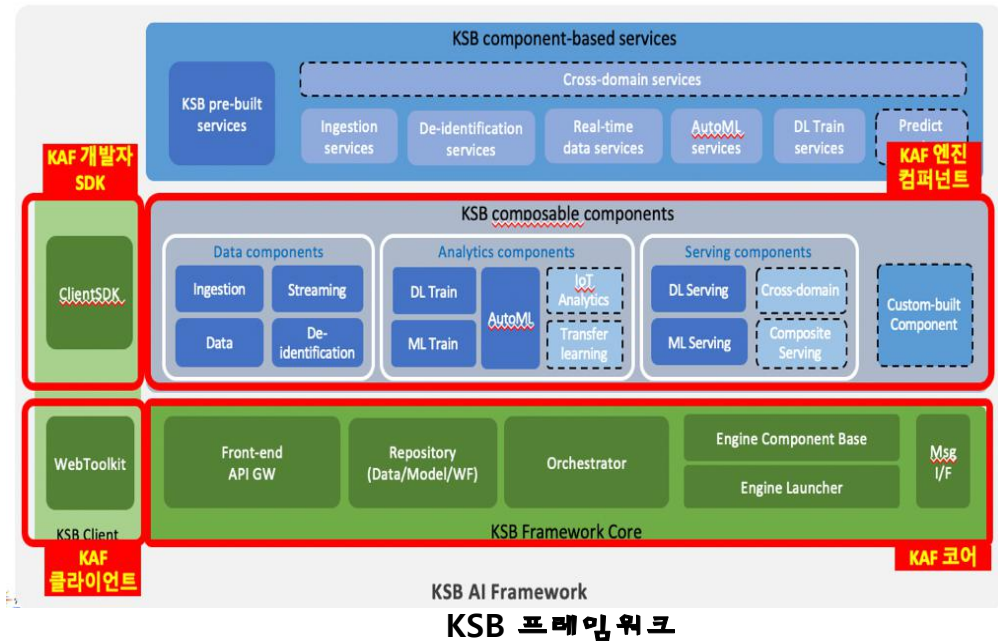
❖ 다양한 도메인의 인공지능 서비스를 생성하고 제공하기 위해 필요한 공통 기능을 제공하는 프레임워크 기술로서 다음을 포함

- KAF 코어 모듈: 엔진을 구성하고 실행하는 프레임워크 코어
- KAF 컴포넌트 모듈: 서비스 제공에 있어 필요한 기능들을 미리 구현해놓은 컴포넌트
- KAF 클라이언트 모듈: 워크플로우를 정의할 수 있는 웹저작도구
- KAF 개발자 SDK 모듈: 엔진 컴퍼넌트 및 개발환경(오픈소스 정책)

기술 구성도

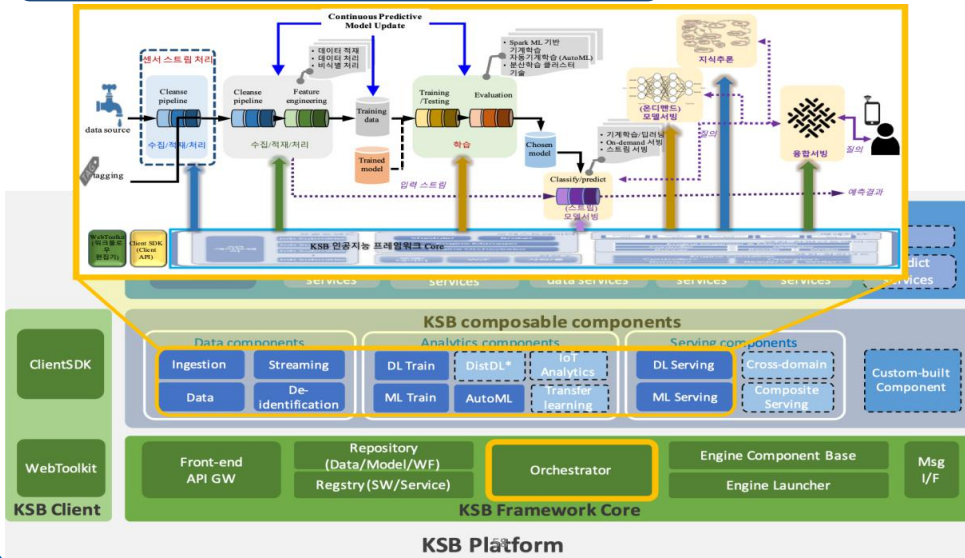


KSB Client (웹툴킷)



1. 기술의 개요 - 세부 기능 설명

KAF 코어 모듈

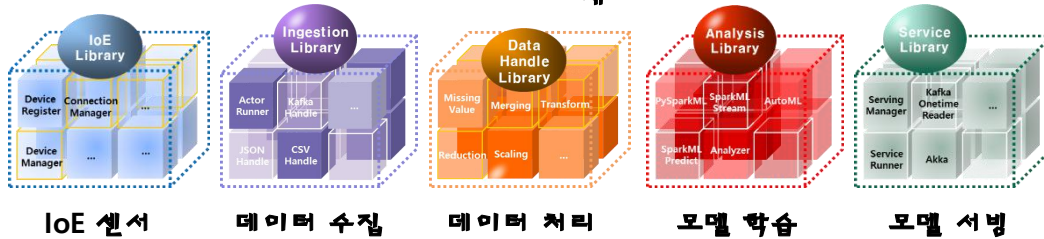


주요 기능

- 다양한 목적의 서비스를 제공하는 엔진을 구성하는 KAF엔진 컴퍼넌트 유형들의 규격 제공
- 규격을 만족하는 하나 이상의 KAF엔진 컴퍼넌트들을 조합해서 동적으로 엔진 구성
- 워크플로우 실행 요청에 대해 하나 이상의 엔진 인스턴스를 생성
- 워크플로우 내의 일련의 엔진들에 대한 스케줄링 기능을 제공
- 실행 중인 워크플로우를 제어

KAF 컴포넌트 모듈

인공지능 서비스를 위한 필요 기능들을 미리 탑재

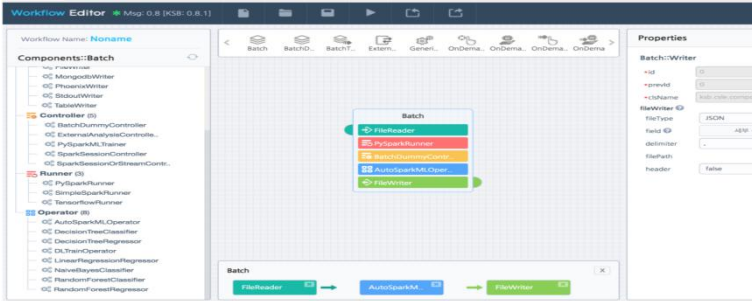


주요 기능

- 다양한 저장소로부터 데이터 가져와 적재
- 데이터 배치/스트림 처리
- 머신러닝/딥러닝 모델 학습 기능
- 자동기계학습 기능
- 머신러닝/딥러닝 모델을 스트림/온디멘드 방식으로 서빙
- 서빙 엔진의 실행에 의해 생성된 하나 이상의 서비스를 파이프라인하여 융합 서비스 제공

1. 기술의 개요 - 세부 기능 설명

KAF 클라이언트 (웹저작도구 포함)



■ 주요 기능

- 사용자가 컴포넌트를 조합하여 워크플로우를 구성할 수 있는 GUI 환경의 저작 기능을 제공
- 저장소의 데이터에 대한 시각화 기능을 제공
- 워크플로우의 실행 상태를 모니터링할 수 있는 GUI 환경 제공
- 사용자가 워크플로우를 제어할 수 있는 클라이언트 도구 제공
- 사용자 관리 기능
- 신규 엔진 컴퍼넌트를 저작도구에서 사용할 수 있도록 등록 및 관리할 수 있는 GUI 환경 제공

KAF 개발자 SDK

Operator 컴퍼넌트 개발하기

컴퍼넌트를 Protocol Buffer 메시지로 작성
 새로운 프레임워크 코어 노트하기
 새로운 프레임워크 코어 프로젝트 만들기
 신규 컴퍼넌트 소스라 코드 작성
 신규 컴퍼넌트 빌드하여 JAR 파일 생성하기
 신규 컴퍼넌트를 실행 프레임워크에 배포

Operator 컴퍼넌트 개발하기

데이터 처리용 Operator 컴퍼넌트를 만들기 위한 절차는 다음과 같습니다.

- 신규 컴퍼넌트를 Protocol Buffer 메시지로 작성
- 새로 작성한 메시지를 포함한 프레임워크 코어 새로 빌드
- Operator 컴퍼넌트 소스라 코드 작성
- 컴퍼넌트 빌드하여 프레임워크에 배포

아래에는 Operator 컴퍼넌트 개발하고 이를 실행 중인 KSB 프레임워크 상에 배포하는 과정을 설명합니다. 자세한 설명을 위해서 column selection을 수행하는 간단한 오피레이팅인 ColumnSelectWithFileOperator 개발을 대상으로 설명을 하도록 하였습니다. 단, 이미 ColumnSelectWithFileOperator이 등록되어 있으므로, 여기서는 MyColumnSelectWithFileOperator로 변경해서 신규 컴퍼넌트로 작성하여 등록하는 과정을 진행해 보도록 하였습니다.

컴퍼넌트용 Protocol Buffer 메시지 작성

우선, 개발할 MyColumnSelectWithFileOperator에 전달될 파라미터를 Google의 Protocol buffer 메시지로 작성해야 합니다. 이를 위해서 다운로드한 KSB 프레임워크 오픈소스용 ScalaIDE를 이용하여 프로젝트들 옆에 ksb-core-digest_2.11 프로젝트가 있는지 확인합니다.

ksb-core-digest_2.11 모듈은 엔진을 생성하고 구동하기 위한 컴퍼넌트들의 파라미터를 정의한 Protocol buffer 메시지를 모아 놓은 모듈이며, 다음과 같은 구조를 갖습니다.

■ 주요 기능

- 신규 엔진 컴퍼넌트 개발 환경 제공
- 신규 개발 엔진 컴퍼넌트를 이용한 워크플로우 개발 환경 제공
- 개발한 워크플로우의 실행을 KAF 코어에 요청할 수 환경 제공

2. 기술미전 내용 및 범위

보급형 KSB 인공지능 프레임워크 기술 (KAF)

□ 기술미전 내용

구분	기술모듈	분류	주요기능
1	KAF 코어	프레임워크 기능	<ul style="list-style-type: none"> 다양한 목적의 서비스를 제공하는 엔진을 구성하는 KAF엔진 컴퍼넌트 유형들의 규격 제공 KAF엔진 컴퍼넌트들을 조합해서 등적으로 엔진을 구성하는 기능 구성된 엔진들을 실행요청에 의해 엔진 인스턴스를 생성하는 기능
		1KAF 코어 플랫폼 기능	<ul style="list-style-type: none"> 엔진의 실행을 스케줄링하는 기능 각 엔진의 실행을 구동하는 기능 실행 중인 엔진을 제어하는 기능
2	KAF 엔진 컴퍼넌트	데이터 적재 기능	<ul style="list-style-type: none"> 다양한 저장소로부터 데이터를 가져와 다른 저장소로 적재하는 기능 스트림 데이터를 적재하는 기능
		2KAF 엔진 컴퍼넌트 데이터 처리 기능	<ul style="list-style-type: none"> 데이터를 배치 처리하는 기능 스트림 데이터를 처리하는 기능
		2KAF 엔진 컴퍼넌트 모델 학습 기능	<ul style="list-style-type: none"> 머신러닝 모델 학습 기능 딥러닝 모델 학습 기능 자동머신러닝 (AutoML) 모델 학습 기능
		2KAF 엔진 컴퍼넌트 학습 모델 서빙 기능	<ul style="list-style-type: none"> 머신러닝 모델의 스트림 방식 서빙 기능 딥러닝 모델의 스트림 방식 서빙 기능 딥러닝 모델의 On-Demand 방식의 서빙 기능
		2KAF 엔진 컴퍼넌트 통합서빙 기능	<ul style="list-style-type: none"> 서빙에 의해 생성된 하나 이상의 서비스를 파이프라인하여 새로운 서빙을 제공하는 기능
3	KAF 클라이언트	워크플로우 저작 기능	<ul style="list-style-type: none"> 워크플로우를 DIY 방식으로 정의할 수 있는 저작 기능 데이터의 시각화 기능
		3KAF 클라이언트 워크플로우 실행 모니터링 및 제어 클라이언트 기능	<ul style="list-style-type: none"> 워크플로우 단위의 실행상태를 모니터링 기능 엔진 단위의 실행상태 모니터링 기능 워크플로우 단위의 실행 제어를 요청하는 기능 엔진 단위의 실행 제어를 요청하는 기능
		3KAF 클라이언트 관리 기능	<ul style="list-style-type: none"> 사용자 관리 기능 신규 엔진 컴퍼넌트의 저작도구 등록 및 관리 기능 엔진 컴퍼넌트셋의 리비전 기능
4	KAF 개발자 SDK	신규 컴퍼넌트 개발 기능	<ul style="list-style-type: none"> 신규 엔진 컴퍼넌트 개발 환경 제공 워크플로우 실행 기능

□ 기술미전 범위

- ❖ KAF 코어 SW (바미너리 제공)
- ❖ KAF 클라이언트 SW (바미너리 제공)
- ❖ 요구사항 및 구조설계 문서
- ❖ 설치 및 사용자 매뉴얼
- ❖ KAF 엔진 컴퍼넌트 SW (오픈소스 라이선스 정책)
- ❖ KAF 개발자 SDK SW (오픈소스 라이선스 정책)

2. 기술이전 내용 및 범위

□ 기술 개발 현황

❖ 기술성숙도(TRL : Technology Readiness Level) 단계 : (7)

구분	단계	정 의	세 부 설 명
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	◦ 시험샘플을 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 생산물에 대해 생산량, 생산용량, 불량률 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

KSB 인공지능 프레임워크

▣ 경쟁기술/대체기술 현황

- ❖ Google은 Cloud 상에 머신러닝 라이프사이클을 구현한 Kubeflow Pipeline 상용 서비스 시작
- ❖ Databricks는 딥러닝 포함한 다양한 머신러닝 플랫폼과의 파이프라인을 제공하는 Azure Databricks 서비스 제공
- ❖ GE는 자신들의 IoT 경험을 바탕으로 클라우드 기반의 데이터 분석 기술을 통합한 산업형 IoT 플랫폼 출시

▣ 주요특징

- ❖ 산업형 인공지능 솔루션 개발 측면에서 단순 머신러닝 파이프라인 기술이나 사물인터넷 플랫폼과 차별화할 수 있는 산업형 인공지능 전주기 솔루션기술을 확보할 수 있음
- ❖ 데이터 수집부터 데이터 처리, 모델 관리, 모델 (융합) 서빙까지 인공지능 서비스 관련 전주기적 솔루션 제공
- ❖ DIY 기반 도메인 맞춤형 AI 서비스 구축 용이

4. 기술의 사업성 - 예상 제품 및 서비스

▣ 예상 응용 제품 및 서비스

활용 분야	제품 및 활용분야 세부 내용
KSB 인공지능 프레임워크 어플라이언스	<ul style="list-style-type: none"> • 지능형 데이터 저장소 어플라이언스 • 지능형 IoT 플랫폼 어플라이언스
AI 기반 산업 솔루션	<p>데이터 수집부터, 학습 모델 개발, 배포, 스케일아웃을 포함한 인공지능 서비스 전체 라이프 사이클을 빠르게 구현 및 도메인 특화 인공지능 서비스 프로덕션화 가능</p> <ul style="list-style-type: none"> • 재실자 편의성 제공 및 빌딩에너지 사용량 최적화를 위한 스마트 에너지 수요공급 시스템 • 실시간 교통 관제 및 예측을 통한 차량 운행 관리 시스템 • 건강 정보 관리를 통한 개인 맞춤형 질병 예측 및 예방 시스템 • 시설물/설비 고장 진단 및 예측을 통한 플랜트 안전 시스템 • 악성코드 분석 및 프로파일링을 통한 비정상 행위 탐지 및 대응 시스템
KSB 인공지능 프레임워크 기반 빅데이터 및 인공지능 교육 플랫폼	<p>사용자 친화형 인공지능 교육 서비스 제공</p> <ul style="list-style-type: none"> • GUI를 통해 데이터 수집부터 학습 모델 서빙까지 손쉽게 빠르게 Production화 할 수 있는 교육 프로그램 제공 • 빅데이터, AI 분야 고급인력 확보를 위한 정부 정책에 부합

4. 기술의 사업성 - 시장 규모 및 예상 매출액



▣ 해당 제품/서비스 시장 규모

관련 제품/서비스 시장	시장	관련 시장 규모 (5년), (단위: 백만불, 십억원)				
		2020	2021	2022	2023	2024
KSB 인공지능 프레임워크 어플라이언스	해외	338.49	518.01	779.85	1139.61	1596.93
	국내	273	333	399.6	479.52	575.4
AI 기반 산업 솔루션	해외	676.98	1036.02	1559.7	2279.22	3193.86
	국내	546	666	799.2	959.04	1150.8
KSB 기반 빅데이터 및 인공지능 교육 플랫폼	해외	112.83	172.67	259.95	379.87	532.31
	국내	91	111	133.2	159.84	191.8
합계	해외	1128.3	1726.7	2599.5	3798.7	5323.1
	국내	910	1110	1332	1598.4	1918

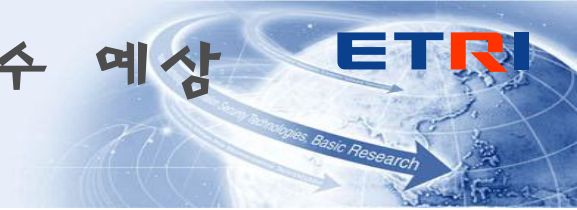
※ 전체 시장은 Statistica 2018년 자료, 2017년 미래창조과학부 자료를 참조
 ※ 관련 제품/서비스 시장은 전체 인공지능 시장의 10%로 산정 (x)
 ※ 일반적인 인공지능 어플라이언스 시장은 x의 30%로 산정, 도메인 특화 AI 산업 솔루션 시장은 x의 60%로 산정, 인공지능 전문인력 양성 수요 급증으로 x의 10%를 교육 플랫폼 관련 시장으로 산정함

▣ 예상 매출액

관련 제품/서비스 시장	시장	관련 시장 규모 (5년), (단위: 백만불, 십억원)				
		2020	2021	2022	2023	2024
KSB 인공지능 프레임워크 어플라이언스	해외	3.38	6.48	12.18	22.26	38.99
	국내	13.65	20.81	31.22	46.83	70.24
AI 기반 산업 솔루션	해외	6.77	12.95	24.37	44.52	77.98
	국내	27.3	41.63	62.44	93.66	140.48
KSB 기반 빅데이터 및 인공지능 교육 플랫폼	해외	1.13	2.16	4.06	7.42	12.99
	국내	4.55	6.94	10.41	15.61	23.41
합계	해외	11.28	21.58	40.62	74.19	129.95
	국내	45.5	69.38	104.06	156.09	234.13

※ 국내 AI 분야 기술수준을 고려(최고수준 보유국 미국대비 78.86%, 격차 1.71년, ICT 기술수준조사보고서, 2018.2, IITP)하여 세계 시장의 1%점유 후 CA GR 25%로 증가한다고 가정
 ※ 1단계 연구 결과물인 KSB 인공지능 프레임워크를 활용하여 인공지능 서비스를 제공하고자 하는 기업의 수는 현재 29개이며 점차 늘어날 예정임. 2019년 국내 시장의 5% 점유 후 CAGR 25%로 증가한다고 가정

4. 기술의 사업성 - 매출정률사용료 징수 예상액



▣ 매출정률 사용료 징수 예상액

단위: 억원

구분	예상매출액 (5년간)	ETRI 기술 차지 비중	매출정률	매출정률 사용 예상액
대기업	914.56	20%	5%	9.146
중견기업	914.56	20%	3.75%	6.859
중소기업	7316.49	20%	1.25%	18.291

※ 1달러 = 1100원으로 산정

※ KSB 프레임워크를 활용하여 인공지능 서비스를 제공하고자 하는 기업 중 대부분이 중소기업임을 감안하여, 전체 매출액의 80%는 중소기업이, 10%는 대기업이, 그리고 10%는 중견기업이 달성한다고 가정

※ KSB 인공지능 프레임워크 상에 중소기업 자체 솔루션을 추가 개발하여 사업화할 것을 고려, 수익창출 또는 비용절감에 공헌한 ETRI 기술 차지 비중을 예상 매출액의 20%로 산정

5. 국내외 시장 동향

□ 국외 기술현황

- ❖ 최근 머신러닝을 위한 개발부터, 배포, 스케일아웃을 포함한 ML 워크플로우 라이프사이클을 오케스트레이션할 수 있는 다양한 기술들이 오픈소스로 공개됨. 또한 머신러닝을 현장에 활용하기 위한 엣지-클라우드를 연동한 상용 플랫폼 서비스 시작
- ❖ Kubeflow는 Kubernetes를 이용한 마이크로서비스 기반의 오픈소스 플랫폼으로서 다양한 머신러닝 플랫폼으로 확장가능한 구조로서 ML모델의 개발과 서빙배포에 특화되어 있으며 시각화 제공. Apache 2.0 라이선스로서 아직은 초기 단계이나 구글의 지원을 받으며 현재 120명의 커미터가 활동중임
- ❖ 또한, 구글은 Google Cloud 상에 머신러닝 라이프사이클을 파이프라인 기술 구현한 Kubeflow Pipeline의 상용 서비스를 시작함
- ❖ Spark 플랫폼 개발을 주도하고 있는 DataBricks는 Spark 스트림처리와 Spark MLlib을 통합한 SparkML pipeline 기술을 이미 완성, 딥러닝 및 포함한 다양한 머신러닝 플랫폼과의 파이프라인을 제공하는 Azure Databricks 서비스를 Microsoft Azure 클라우드 상에서 서비스하고 있음
- ❖ 한편, 산업 분야에서는 미국 GE 사는 자신들의 IoT 경험을 바탕으로 클라우드 기반의 데이터 분석 기술을 통합한 산업형 IoT 플랫폼인 Predix 플랫폼을 상용으로 출시하고, 화학, 제조, 자동차 등 다양한 산업 도메인에 적용

□ 국내 기술현황

- ❖ 삼성은 ML 파이프라인 도구로서 오픈소스 Kubeflow의 도입을 고려하고 있으며 개발상황을 예의주시하고 있음. 또한, 지능형 IoT를 위한 EdgeX 기반의 지능형 엣지 컴퓨팅을 위한 오픈소스 활동에 참여하고 있음
- ❖ 네이버는 ML Serving을 위한 인프라 기술을 독자적으로 구축하여 검색서비스의 인프라 기술로 활용

5. 국내외 시장 동향

■ 세계시장 전망

- ❖ 인공지능기술은 2017년 48억달러에서 2023년 532억달러 규모로 연평균 49.2%의 고도성장미 예상됨(출처 : Statistica 2018)
- ❖ AI SW 플랫폼 시장은 2017년 23억달러 규모에서 2021년 84억달러 규모로 연평균 약 40%의 고속 성장미 예상됨(출처 : IDC 2017)
- ❖ AI 개인비서(IVA: Intelligent Virtual Assistant) 시장은 AI 분야 중 급속히 확산되고 있는 분야로서, 2017년 25억달러 규모에서 2023년 252억달러 규모로 연평균 약 47%의 고속 성장미 예상됨(출처 : Research and Market2018)

■ 국내시장 전망

- ❖ 국내 시장은 현재 초기단계로서, 2017년 6.4조에서 2023년 19.2조 규모로 연평균 20.1%의 고성장이 예측됨(출처 : '2017년 국가정보화에 관한 연차보고서' 의 과기정보통신부 자료 인용)
- ❖ 2020년 기준 음성인식 및 통번역 시장 4.2조원, 영상처리 및 영상인식 시장 3.5조원, 기타 SW 및 알고리즘 관련 시장 3.4조원으로 예상됨

감사합니다.



www.etri.re.kr

※ 하단의 문의처 소개후, 발표후 개별기술 상담이 가능함을 다시 한 번 안내함

♣ 연락처 : KSB융합연구단, 이연희 실장 (042-860-6676, leeyh@etri.re.kr)