

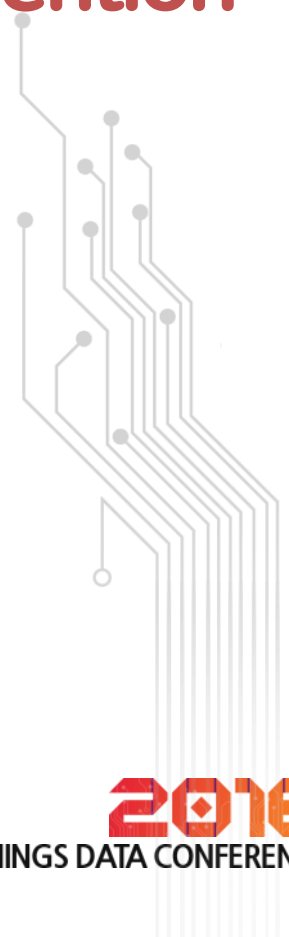


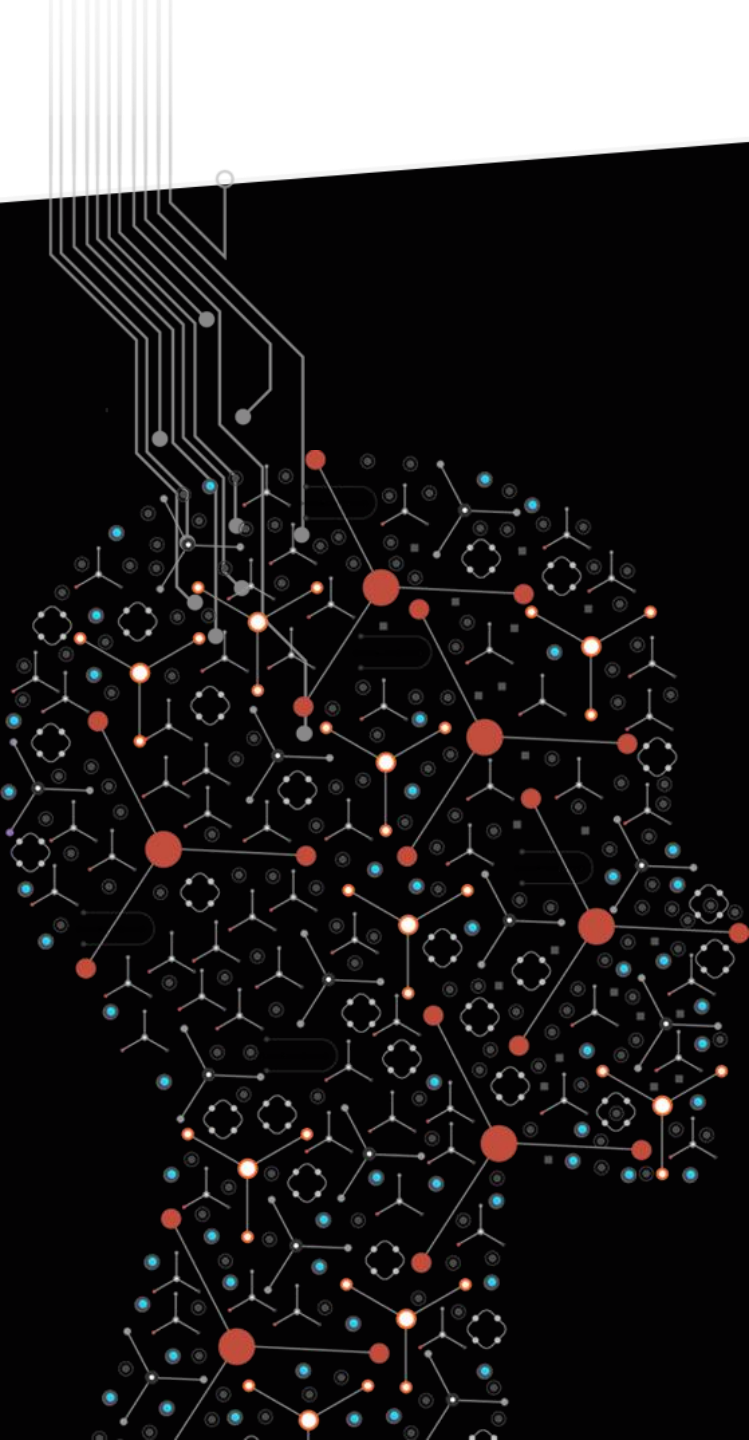
# Data-driven Industry Reinvention

All Things Data Con 2016, Opening speech

SKT 종합기술원

최진성 원장





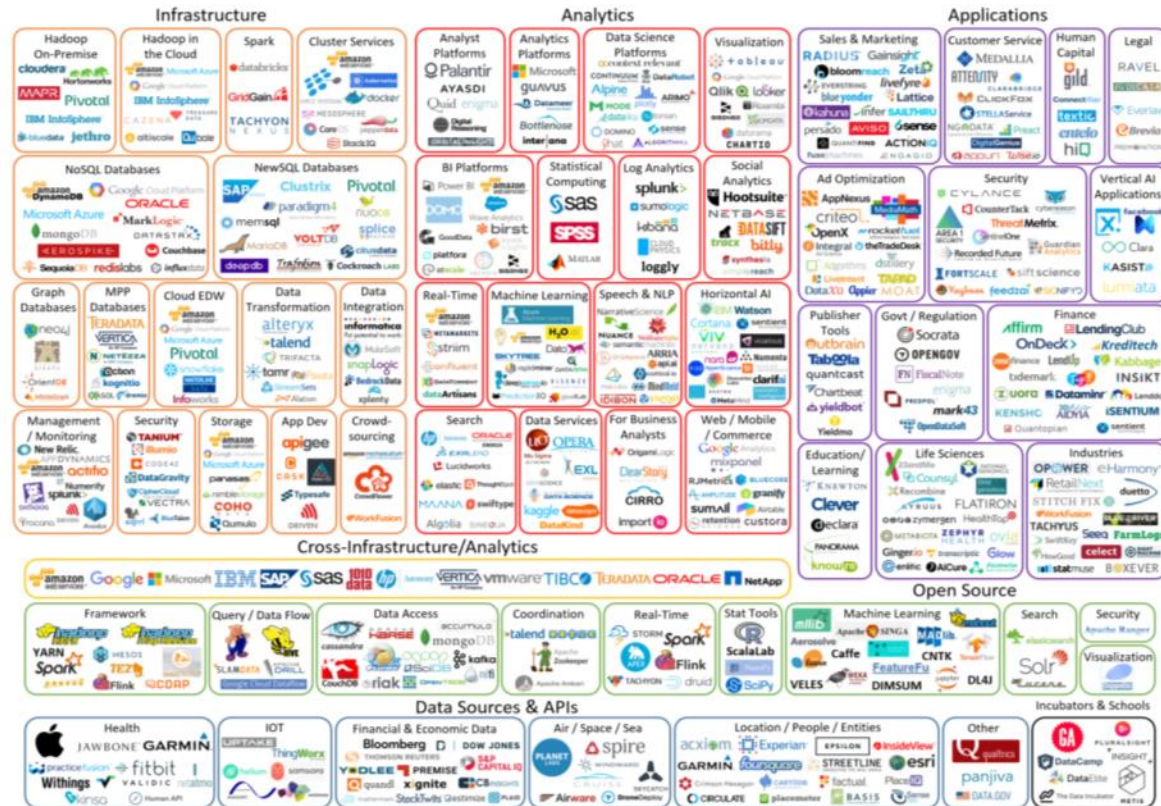
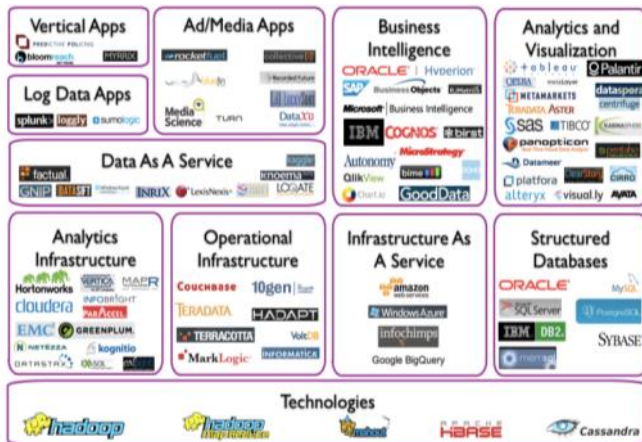
- ❖ Big Data Landscape Expansion
- ❖ Big Data Tech/Biz 진화 방향
- ❖ SK Telecom Big Data Activities
- ❖ Lesson Learned and Other Topics
- ❖ Industry Reinvention

# Big Data Landscape Expansion

Big Data Tech/Biz eco-system은 Intelligence 강화, 다양한 Industry 접목 등을 통해 비약적으로 확장 중

## 2012

## 2016



# 참조. Evolution of Hadoop eco-system

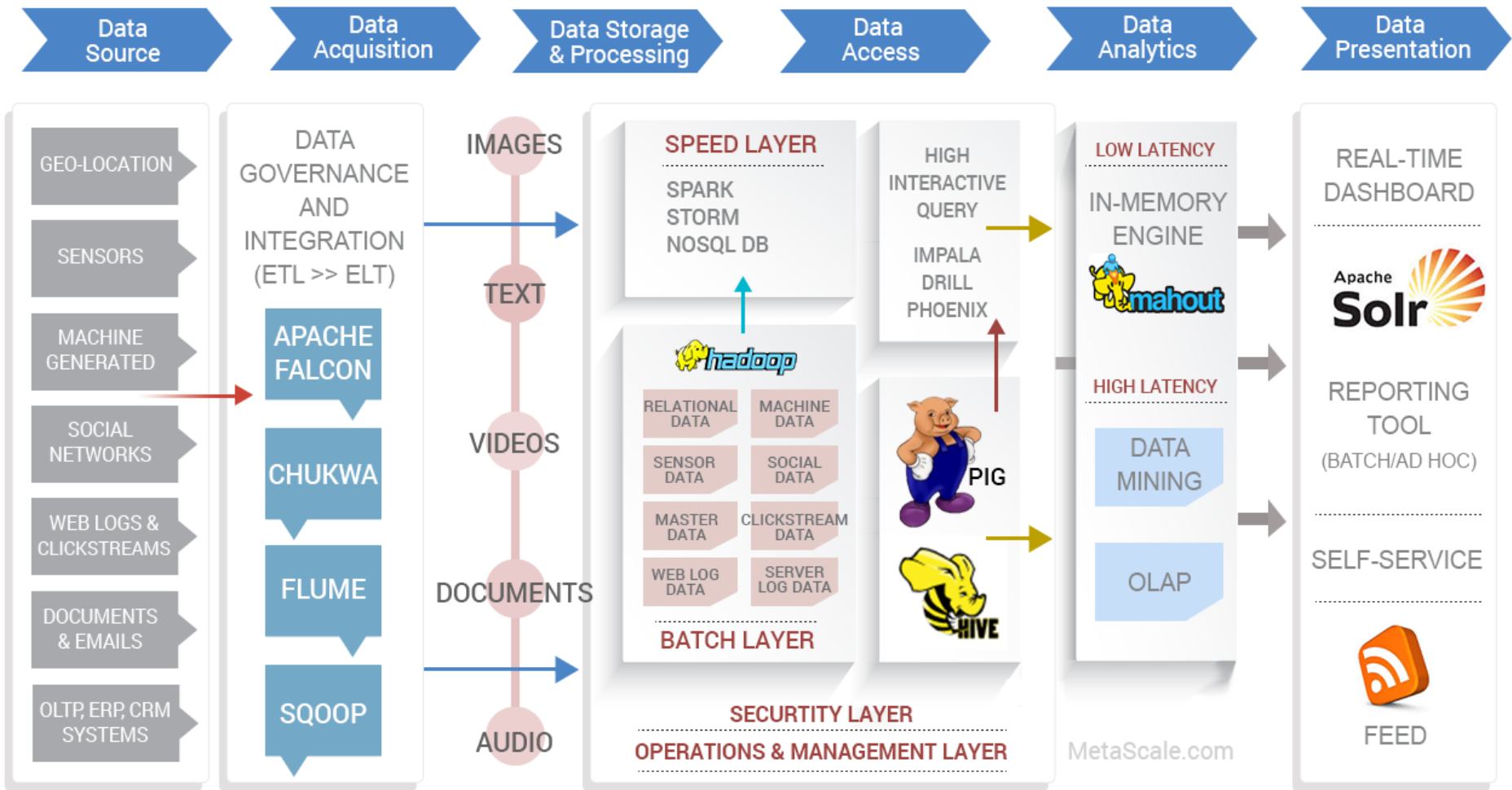
The stack is continually evolving and growing



Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Core Hadoop (HDFS, MR)	Core Hadoop	Core Hadoop	Core Hadoop	Core Hadoop	Core Hadoop	Core Hadoop	Core Hadoop	Core Hadoop	Core Hadoop	Core Hadoop
		Solr Pig	Solr Pig	Solr Pig	Solr Pig	Solr Pig YARN	Solr Pig YARN	Solr Pig YARN	Solr Pig YARN	Solr Pig YARN
		Hbase Zookeeper	HBase ZooKeeper	HBase ZooKeeper	Hive Mahout HBase ZooKeeper	Mahout HBase ZooKeeper	Mahout HBase ZooKeeper	Mahout HBase ZooKeeper	Mahout HBase ZooKeeper	Mahout HBase ZooKeeper
			Hive Mahout	Hive Mahout	Hive Mahout HBase ZooKeeper	Mahout HBase ZooKeeper	Mahout HBase ZooKeeper	Mahout HBase ZooKeeper	Mahout HBase ZooKeeper	Mahout HBase ZooKeeper
				Sqoop Avro	Sqoop Avro	Sqoop Avro	Sqoop Avro	Sqoop Avro	Sqoop Avro	Sqoop Avro
					Flume Bigtop Oozie Hcatalog Hue	Flume Bigtop Oozie HCatalog Hue	Flume Bigtop Oozie HCatalog Hue	Flume Bigtop Oozie HCatalog Hue	Flume Bigtop Oozie HCatalog Hue	Flume Bigtop Oozie HCatalog Hue
						Spark Tez Impala Kafka Drill	Spark Tez Impala Kafka Drill	Spark Tez Impala Kafka Drill	Spark Tez Impala Kafka Drill	Spark Tez Impala Kafka Drill
							Parquet Sentry	Parquet Sentry	Parquet Sentry	Parquet Sentry
								Falcon Knox Flink	Falcon Knox Flink	Falcon Knox Flink
									Kudu RecordService Ibis	Kudu RecordService Ibis



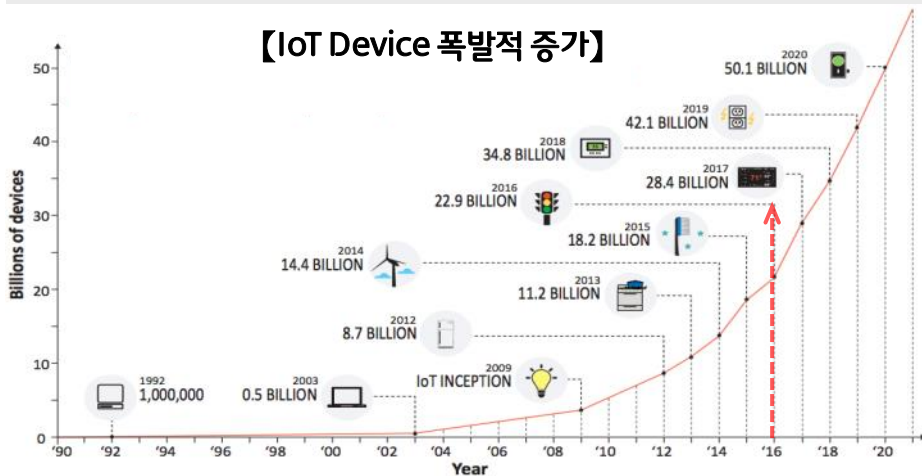
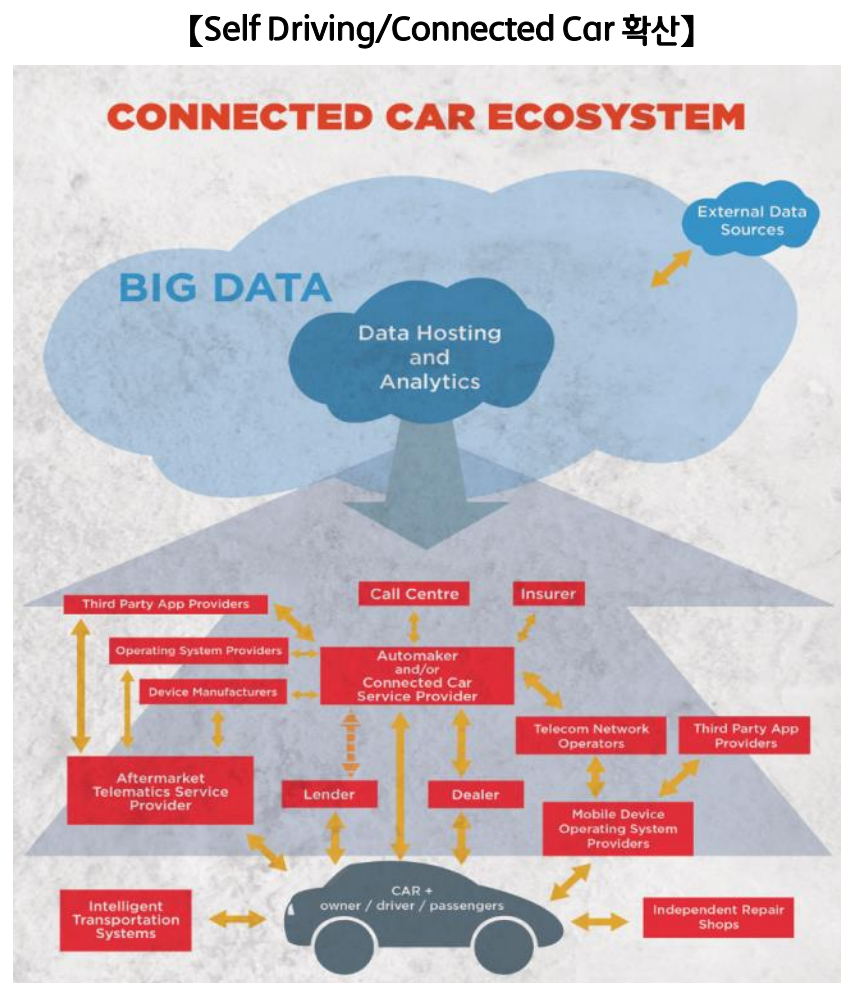
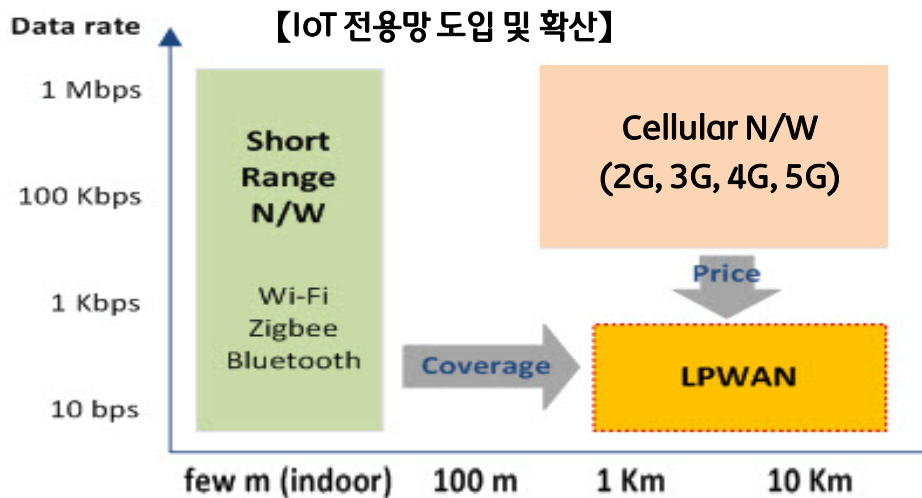
# 참조. Data analysis process



# Big Data Tech/Biz 진화 방향 (1/3)

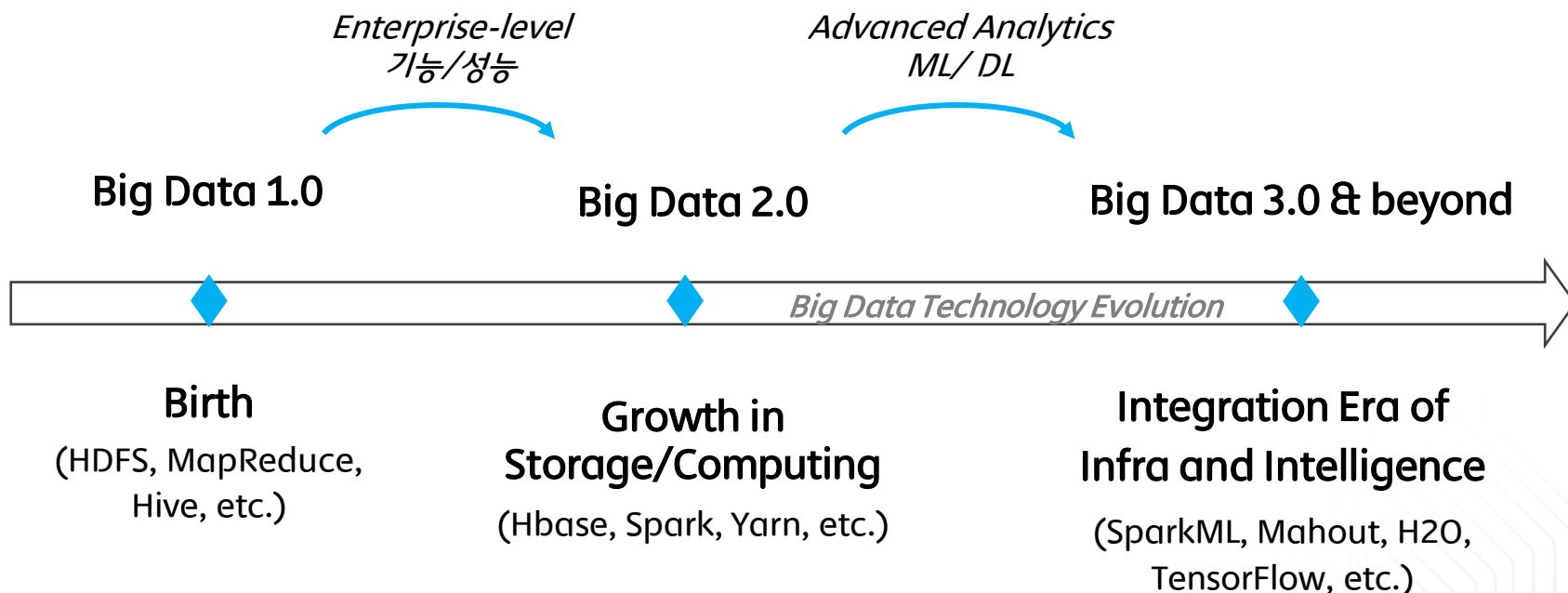
IoT 확산에 따라 심화되고 있는 3V(velocity, volume, variety)에 대응하여 Big Data 기술 진화 중

- **Variety** : Text, web behavior log → Multimedia, machine log
- **Velocity** : Batch → Real Time
- **Volume** : GB, TB → PB, ZB



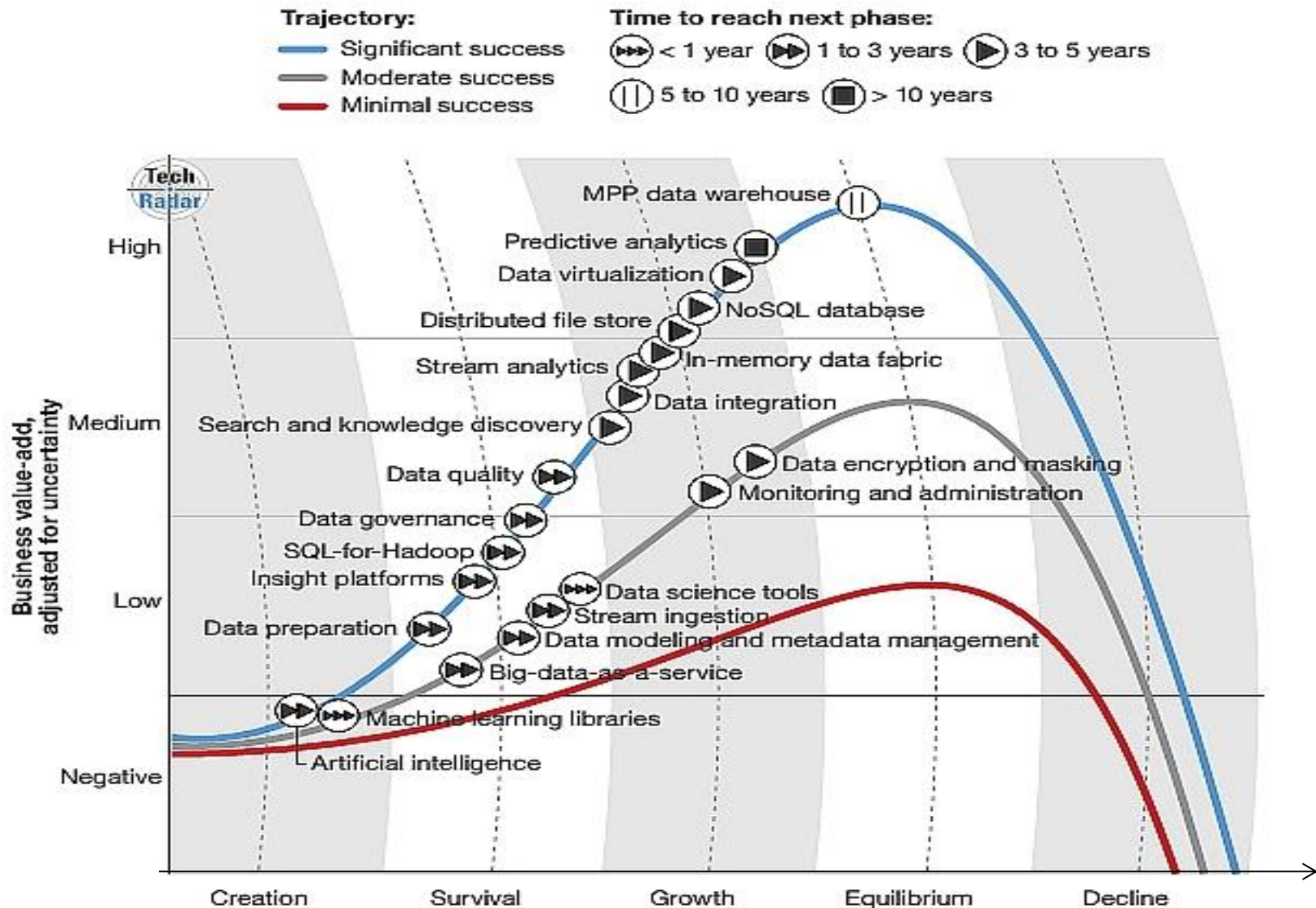
## Big Data Tech/Biz 진화 방향 (2/3)

- Legacy analytics tool을 지원하는 Cost effective Infra (computing/storage) 관점의 활용에서 분산환경 ML/DL 기반의 Intelligence analytics 까지 수용하는 형태로 발전



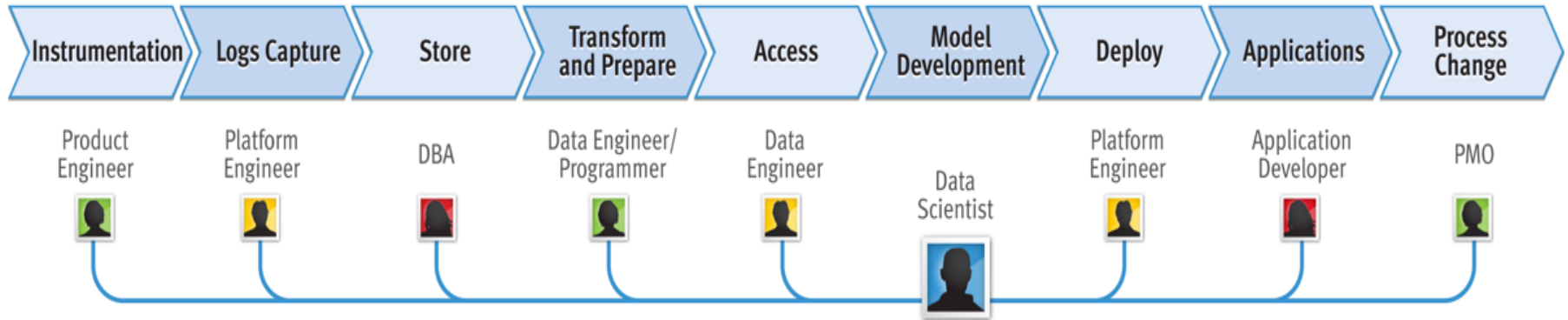
# 참조. Big Data 세부 기술 Life cycle

TechRadar : Big Data, Q1 2016, Forrester Research report








# 참조. Data Science Value Chain



Big Data / AI 기술이 다양한 산업영역에 적용되어 Biz Reinvention 을 일으키고 있음

Various Industry 접목 (As-is)

-  : 항공기 엔진 제조 → 엔진 원격 관리 서비스
- **NOMURA** : AI 기반 주식거래시스템, Robo-adviser
- **Hertz** 5.5조 <  80조  
UBER
-  <  **airbnb**  
24조 35조

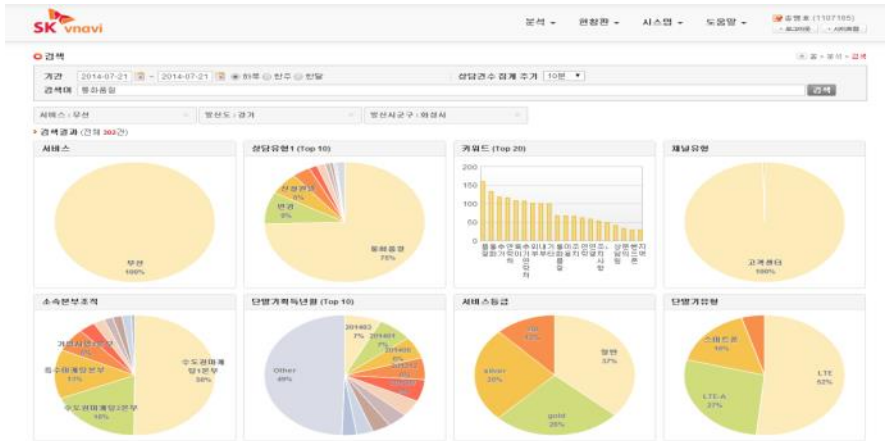
ICT 산업 중심 활용  
(As-was)



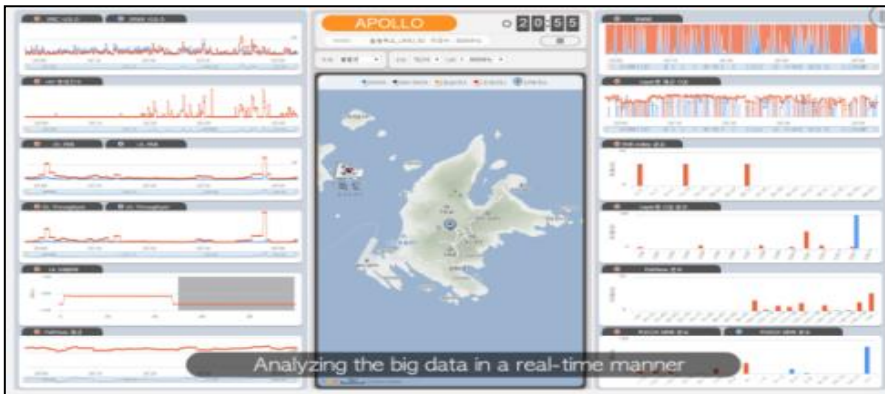
## SKT 내부 활용과 함께 기술 공급자 관점에서 Big Data B2B 솔루션의 외부사업화를 추진 중

### SKT 내부 활용

- Marketing 효율화 : Voc분석, Target marketing 등

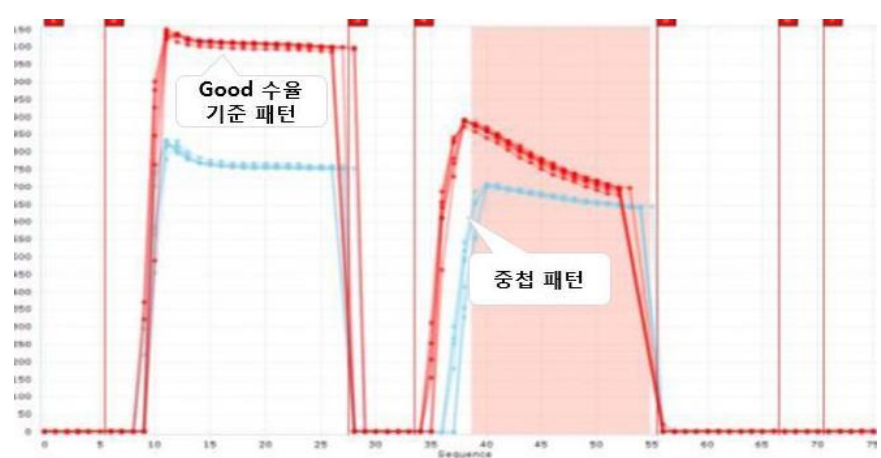


- Network 품질 최적화 : OSS, APOLLO



### 외부 사업화 (Hi-Tech 제조업 등)

- 실시간 FDC (Fault Detection & Classification) 패턴 분석

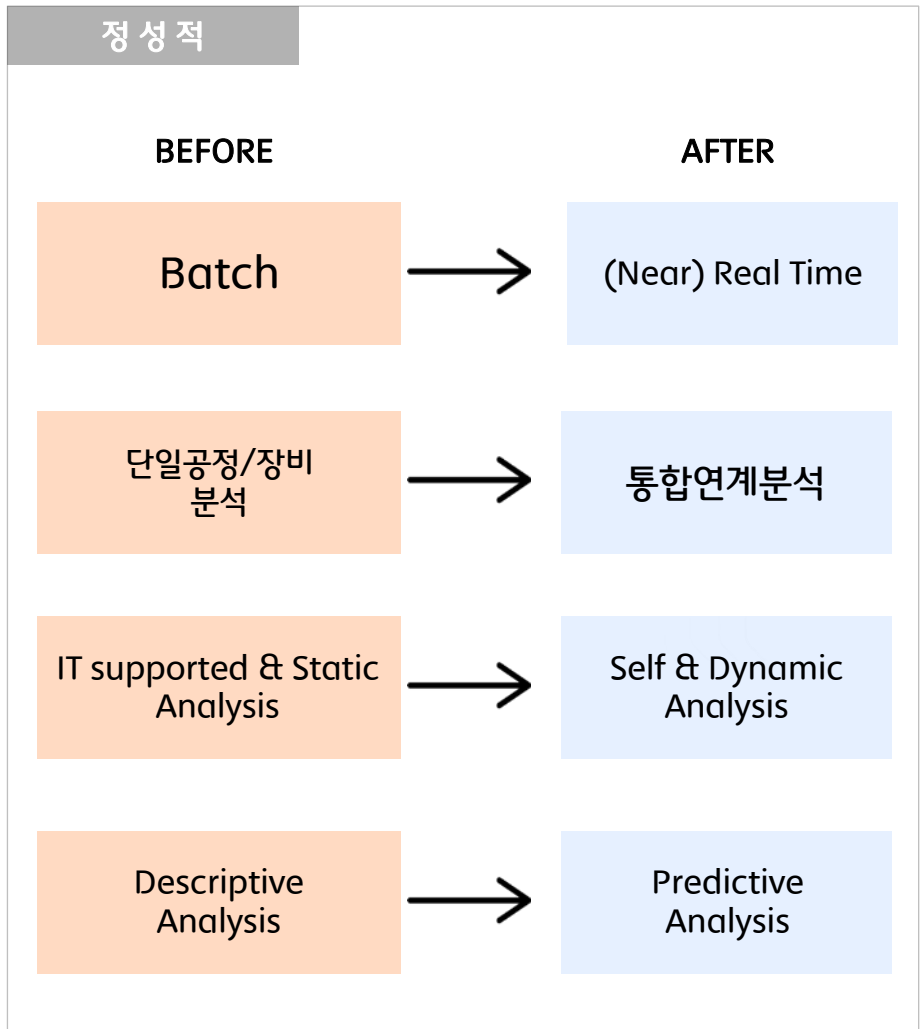
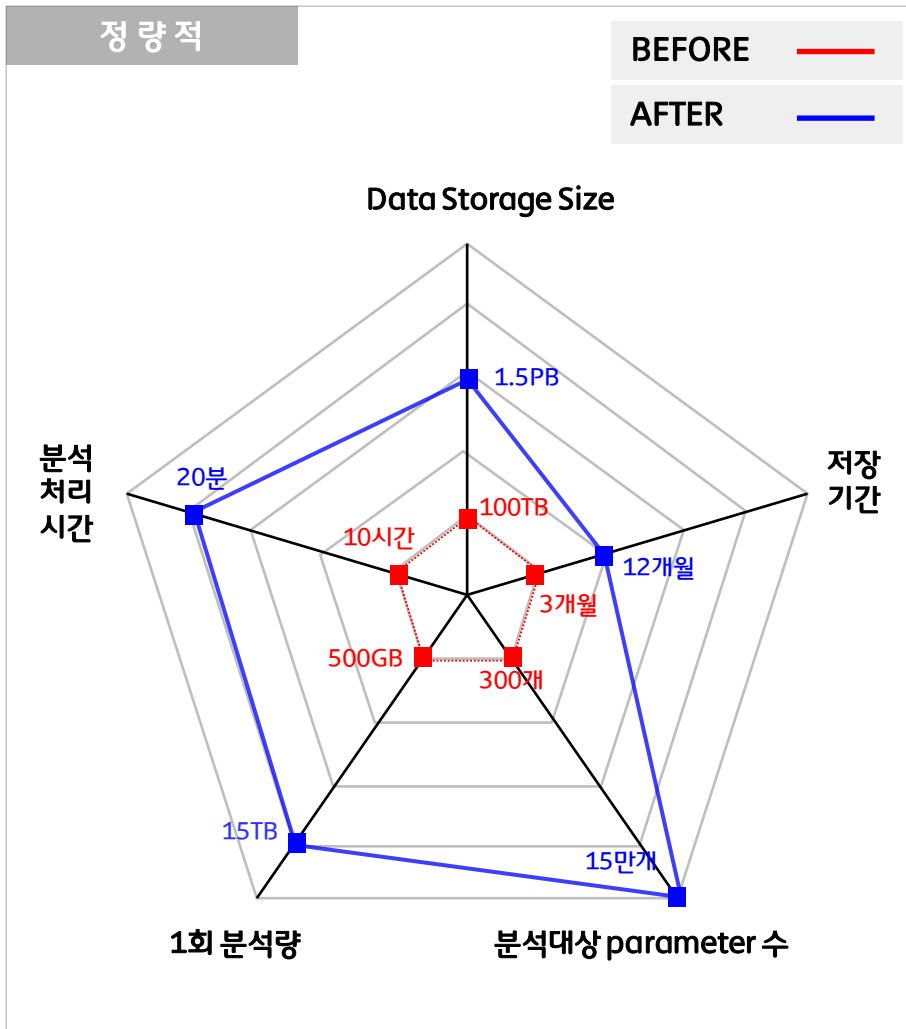


- 대용량 미세 품질 분석



# 참조. Hi-Tech 제조업의 Big data 기술 활용 사례

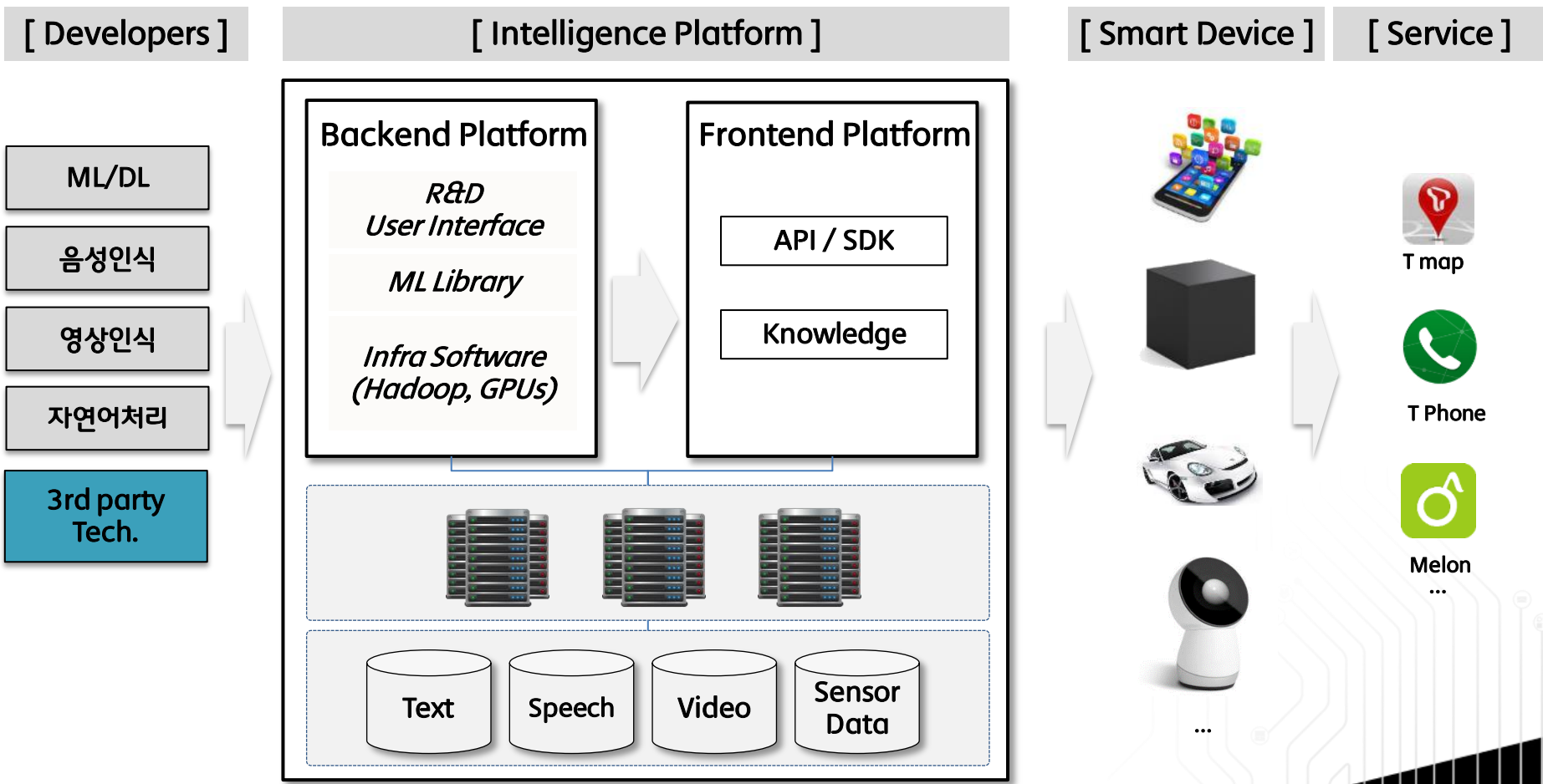
## Big Data 기술 활용 Before & After





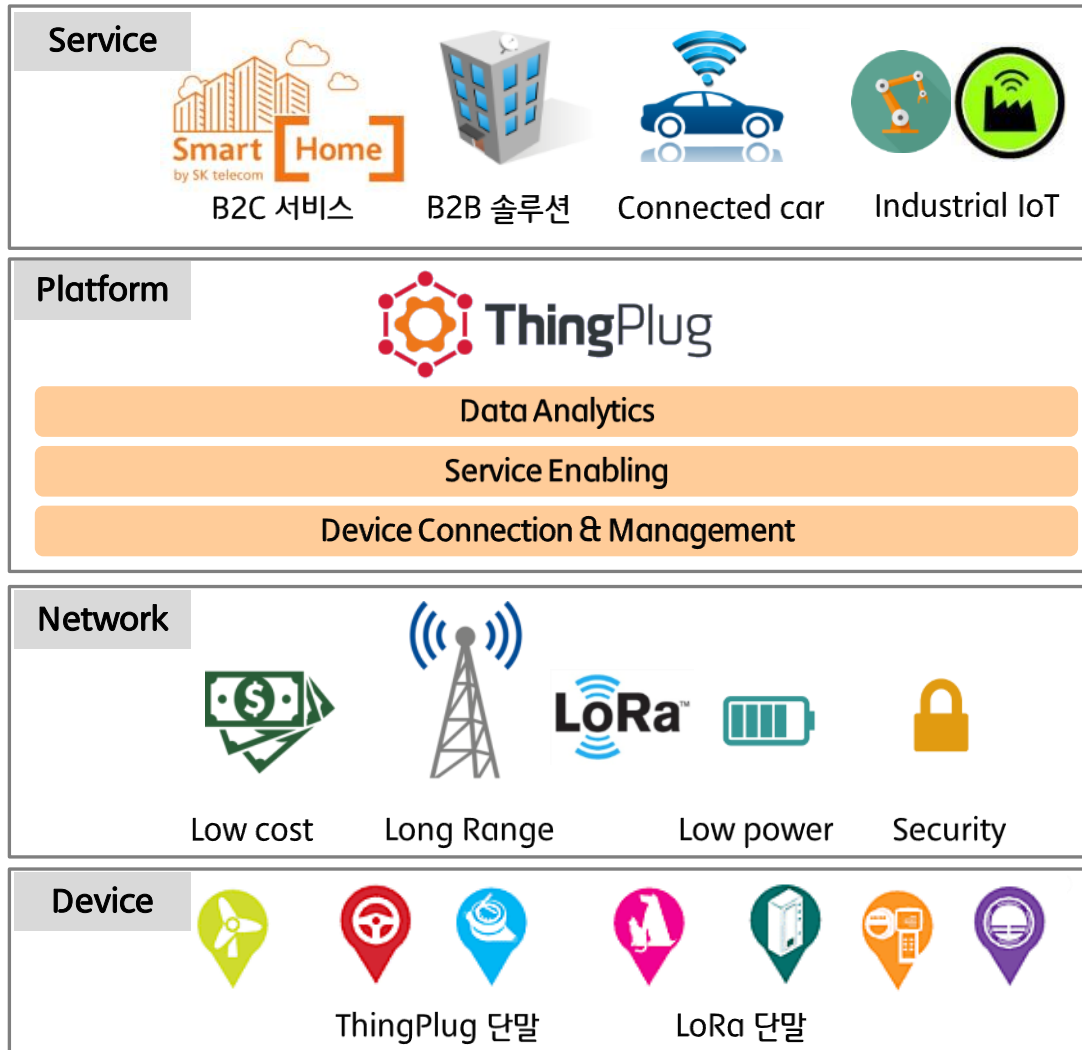
# SK Telecom Big Data 관련 Activities (2/3)

Big Data, 음성/영상인식, 자연어처리 기술 등을 융합하여 intelligence B2C 서비스 제공 추진 중



# SK Telecom Big Data 관련 Activities (3/3)

IoT 확산을 위해 IoT 전용 전국망 (LoRa) 및 IoT platform (ThingPlug) 구축을 추진함





# Lesson Learned and Other Topics

## Lesson Learned

- 시행착오가 필요한 emerging 영역으로서,  
“Small start, Scale later” 접근 방식이 선호됨
- 대부분의 Big data 과제는 BI 에서  
시작하여 New value creation 으로 전이함
- Raw data는 Data lake를 통해 저장하고,  
단위 시스템 별로 필요한 가공 데이터를 feeding  
받는 구조가 이상적인 형태임
- Big data 기술은 다양한 응용/산업과 융합 될 때  
potential 이 더 큰 Enabling tech. 임  
Ex. 자율 주행차, Industry 4.0, O2O 등

## Other Topics

- Machine Learning & Big Data
- In-Memory Processing
- Data Discovery & Self-Service
- LOD & 정부 3.0
- Spark 2.0 & Dataframe, Dataset and  
Structured Streaming
- KeyValue & Distributed File Stores
- Decision support vs. Decision making
- Definition of Intelligent Data
- H2O & Druid

# Industry Reinvention

ICT Tech. 공급자와 Domain 수요자 간의 협업을 통한  
**Industry Reinvention**



Industry Customer

Domain knowledge



Enabling Industry Reinvention



SK텔레콤 & Biz partner

ICT Tech. Expertise

Digitalization



Connectivity



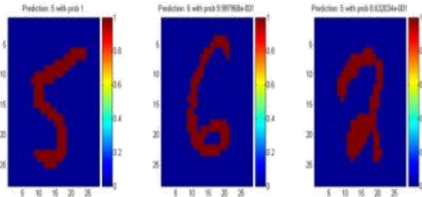
Automation & Intelligence



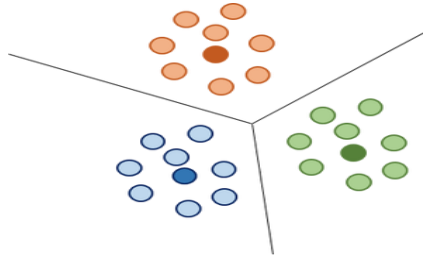


# Appendix 1. Top 10 Machine Learning Algorithms

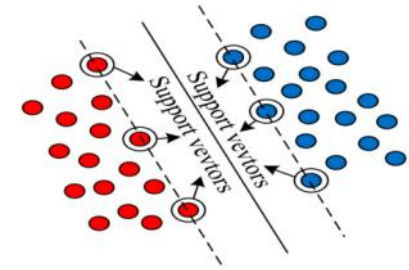
## 1. Naïve Bayes Classifier



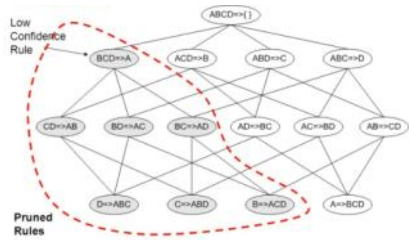
## 2. K-Means Clustering



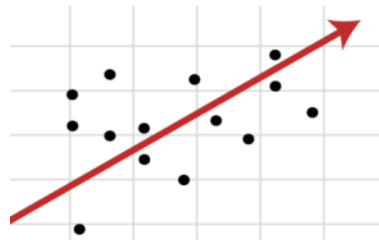
## 3. Support Vector Machine



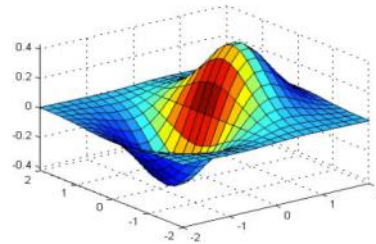
## 4. Apriori Algorithm



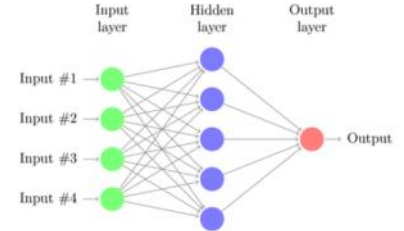
## 5. Linear Regression



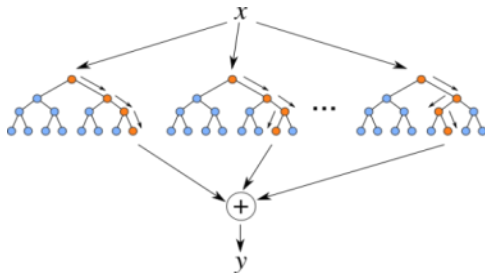
## 6. Logistic Regression



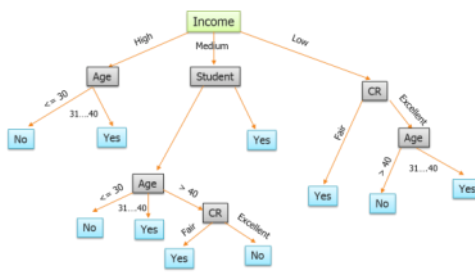
## 7. Artificial Neural Networks



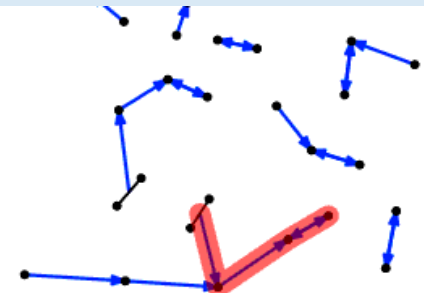
## 8. Random Forests



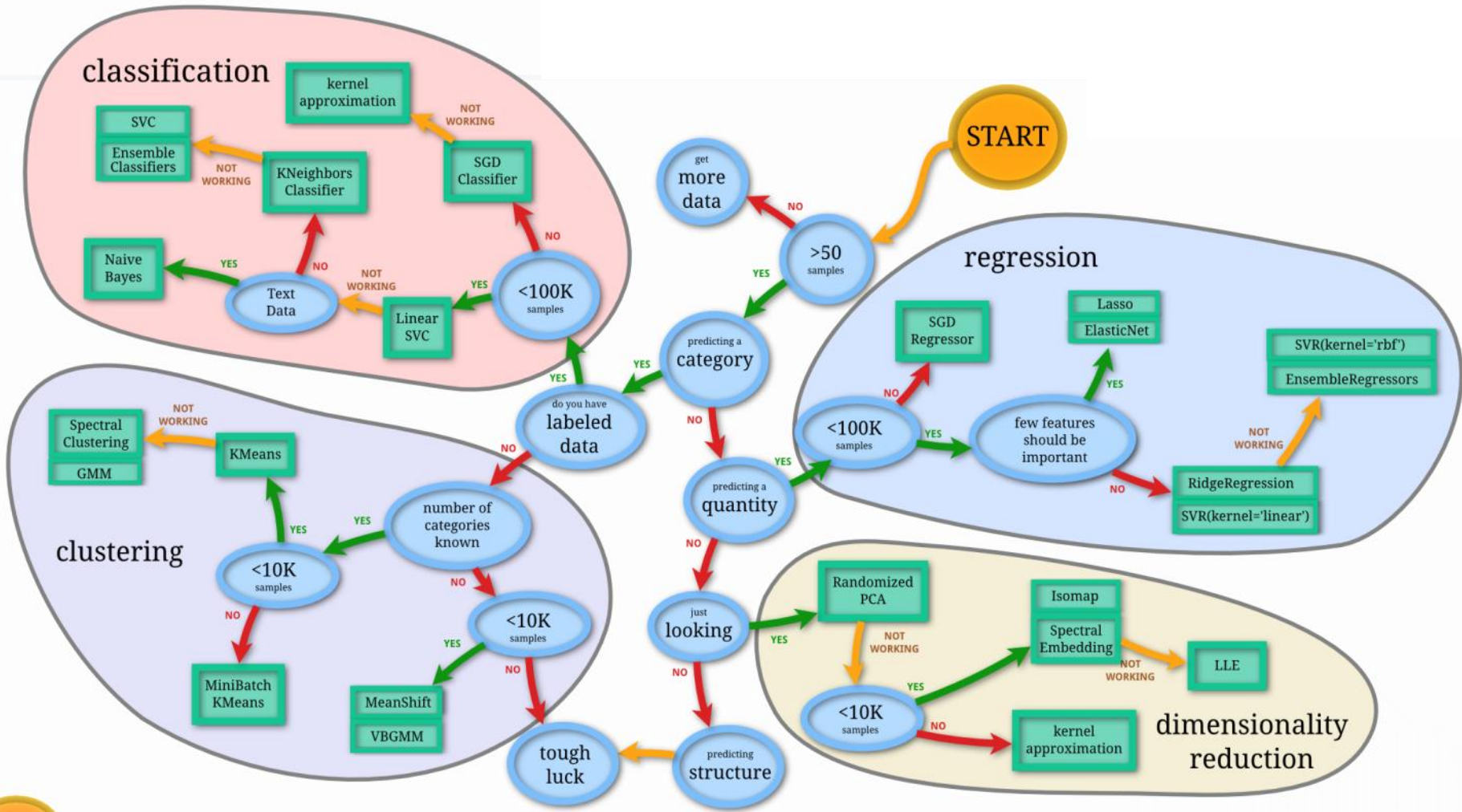
## 9. Decision Trees



## 10. Nearest Neighbors



# Appendix 2. Scikit-learn algorithm cheat-sheet





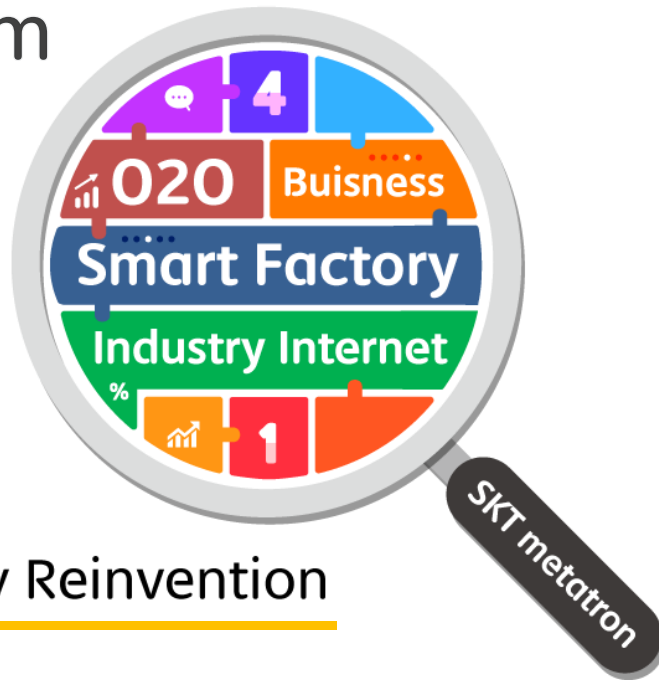
# Appendix 3. Machine Learning Algorithms Mindmap



41회  dev forum

# BIG DATA

Data Driven Industry Reinvention



## Schedule

**7.20 (수)**  
오후 1시 ~ 오후 6시

## Location

**2** 을지로 입구역 4번 출구  
SK T-타워 4층 Supex Hall

## Registration

[developers.sktelecom.com](http://developers.sktelecom.com)  
또는 행사 당일 현장등록 (무료)



# SK Telecom의 Open Innovation channel

SK telecom의 대표 개발자센터인  developers가 새롭게 태어났습니다~!



SK텔레콤의  
다양한 기술 정보



SK텔레콤의 기술을  
활용한 상품 및 서비스 정보



개발자를 위한  
SDK, API 제공



개발자들을 위한  
정보 공유의 장



ICT 기술 동향과 관련된  
새로운 소식



개발자 포럼, 해커톤 등의  
다양한 이벤트 정보



물리적인 개발 환경 및  
인프라 지원

Creative한  
아이디어를  
제안해주세요!

위의 정보를  
한 곳에서 보세요!



 facebook

좋아요를 누르고, 다양한 소식을  
받아보세요!

 YouTube

YouTube를 구독하고,  
최신 기술 소개 및 포럼 영상을 감상하세요!

어떤 디바이스로도 이용이 가능하도록 반응형 웹으로 개발되었습니다.

지금 바로  developers를 경험해보세요~!

<https://developers.sktelecom.com>



*Thank you*

The image features a black background with a white diagonal line running from the top-left to the bottom-right. On the left side, there is a white graphic of a circuit board with several lines extending upwards. On the right side, there is a white graphic of a circuit board with several lines extending downwards. The text "Thank you" is centered in the middle of the page in a white, italicized serif font.