How to Apply Big Data Tools to Policy Making Processes

September 2019

Professor Han, Kyeong Seok, Ph.D.

KOICA

숭실대학교

Professor/Lecturers

Name: Kyeong Seok HAN

Department: Professor of MIS,
College of Business Administration
of Soongsil University
Contact: 010-8954-7137
E-mail: kshan@ssu.ac.kr

<Education Background>

-Professor of MIS, College of Business Administration,
Soongsil University from 1993
-Ph.D., Purdue University, 1989
-Professor of University of Houston, 1990
-Visiting Scholar of The Wharton School, 1999

<Research>

-More than 80 research papers published so far
-Received SFP(Soongsil Fellow Professor) award with the top research performance

Content

I. Can big data revolutionize policymaking by governments?
II. Big Data and E-government
III. Strategies for Korean Big Data Applications
IV. Big Data and Artificial Intelligence for e-Government
V. A Case of Applying Big Data Tools to Strategic Decision Making in a Private Company
VI. Conclusion

I. Can big data revolutionize policymaking by governments?

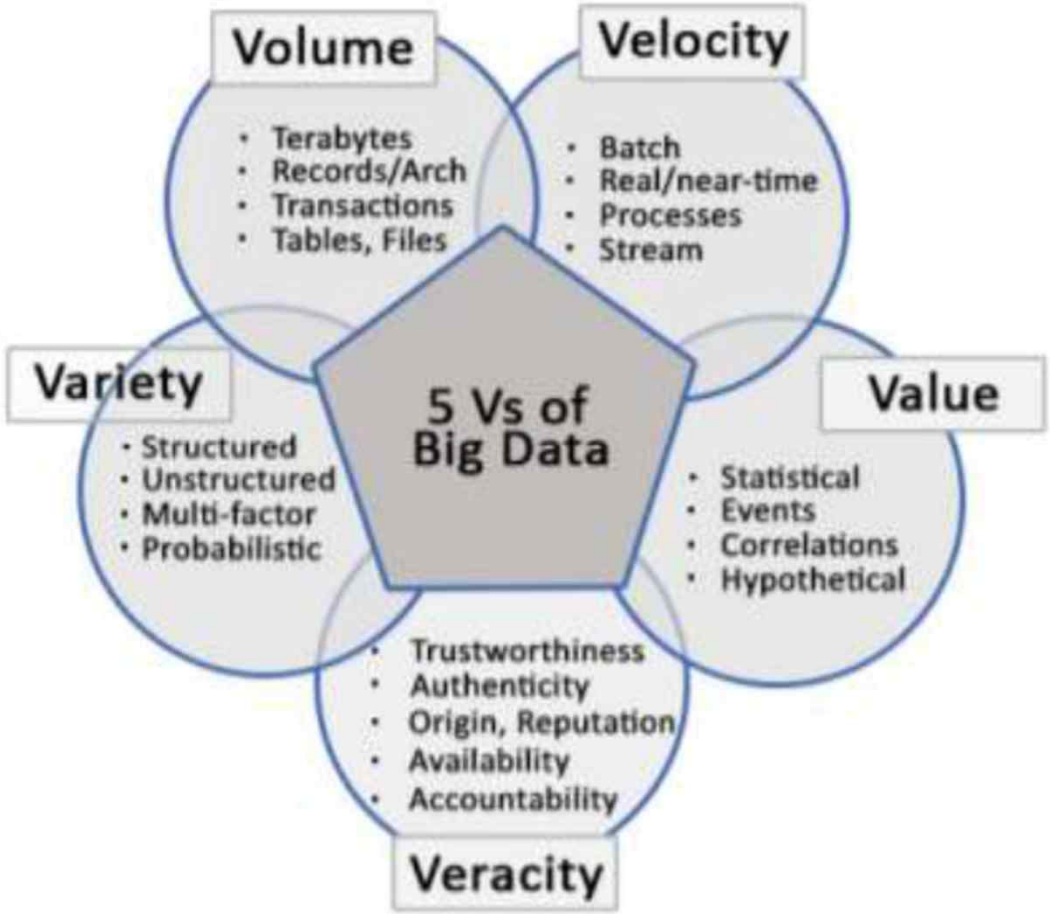
(Source: https://www.ft.com/content/9f0a8838-fa25-11e7-9b32-d7d59aace167, "Can big data revolutionize policymaking by governments?" in Make Big Data open, accessible and available in real time / From Gavin Hayman, Twickenham, Middx, UK Copyright The Financial Times Limited 2018.)

Big Data for Government and the Public Sector

Can big data revolutionize policymaking by governments?

- SAS define big data :
  - "Big data is a popular term used to describe the exponential growth, availability, and use of information, both structured and unstructured."

Can big data revolutionize policymaking by governments?



Text Mining using R

And so even though we face the difficulties of today and tomorrow, I still have a dream. It is a dream deeply rooted in the American dream.

I have a dream that one day this nation will rise up and live out the true meaning of its creed:

We hold these truths to be self-evident, that all men are created equal.

I have a dream that one day on the red hills of Georgia, the sons of former slaves and the sons of former slave owners will be able to sit down together at the table of brotherhood.

I have a dream that one day even the state of Mississippi, a state sweltering with the heat of injustice, sweltering with the heat of oppression, will be transformed into an oasis of freedom and justice.

I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character.

I have a dream today!

I have a dream that one day, down in Alabama, with its vicious racists, with its governor having his lips dripping with the words of interposition and nullification, one day right there in Alabama little black boys and black girls will be able to join hands with little white boys and white girls as sisters and brothers.

Can big data revolutionize policymaking by governments?

I have a dream today!
I have a dream that one day every valley shall be exalted, and every hill and mountain shall be made low, the rough places will be made plain, and the crooked places will be made straight; and the glory of the Lord shall be revealed and all flesh shall see it together.
This is our hope, and this is the faith that I go back to the South with.
With this faith, we will be able to hew out of the mountain of despair a stone of hope. With this faith, we will be able to transform the jangling discords of our nation into a beautiful symphony of brotherhood. With this faith, we will be able to work together, to pray together, to struggle together, to go to jail together, to stand up for freedom together, knowing that we will be free one day.
And this will be the day, this will be the day when all of God s children will be able to sing with new meaning:
My country tis of thee, sweet land of liberty, of thee I sing.
Land where my fathers died, land of the Pilgrim s pride,
From every mountainside, let freedom ring!
And if America is to be a great nation, this must become true.

Can big data revolutionize policymaking by governments?

And so let freedom ring from the prodigious hilltops of New Hampshire.
Let freedom ring from the mighty mountains of New York.
Let freedom ring from the heightening Alleghenies of Pennsylvania.
Let freedom ring from the snow-capped Rockies of Colorado.
Let freedom ring from the curvaceous slopes of California.
But not only that:
Let freedom ring from Stone Mountain of Georgia.
Let freedom ring from Lookout Mountain of Tennessee.
Let freedom ring from every hill and molehill of Mississippi.
From every mountainside, let freedom ring.
And when this happens, when we allow freedom ring, when we let it ring from every village and every hamlet, from every state and every city, we will be able t o speed up that day when all of God s children, black men and white men, Jews and Gentiles, Protestants and Catholics, will be able to join hands and sing in the words of the old Negro spiritual:
Free at last! Free at last!
Thank God Almighty, we are free at last!

Source: http://www.sthda.com/english/wiki/text-mining-and-word-cloud-fundamentals-in-r-5-simple-steps-you-should-know

```
# Install
 install.packages("tm")  # for text mining
 install.packages("SnowballC") # for text stemming
 install.packages("wordcloud") # word-cloud generator
 install.packages("RColorBrewer") # color palettes
 # Load
 library("tm")
 library("SnowballC")
 library("wordcloud")
 library("RColorBrewer")
 # Read the text file from internet
 filePath <- "http://www.sthda.com/sthda/RDoc/example-files/martin-luther-king-i-have-a-dream-speech.txt"
 text <- readLines(filePath)
 docs <- Corpus(VectorSource(text))
 inspect(docs)
 toSpace <- content_transformer(function (x , pattern ) gsub(pattern, " ", x))
 docs <- tm_map(docs, toSpace, "/")
 docs <- tm_map(docs, toSpace, "@")
 docs <- tm_map(docs, toSpace, "₩₩|")
# Convert the text to lower case
docs <- tm_map(docs, content_transformer(tolower))
# Remove numbers
docs <- tm_map(docs, removeNumbers)
```

Can big data revolutionize policymaking by governments?

```
# Remove english common stopwords
docs <- tm_map(docs, removeWords, stopwords("english"))
# Remove your own stop word
# specify your stopwords as a character vector
docs <- tm_map(docs, removeWords, c("blabla1", "blabla2"))
# Remove punctuations
docs <- tm_map(docs, removePunctuation)
# Eliminate extra white spaces
docs <- tm_map(docs, stripWhitespace)
# Text stemming
# docs <- tm_map(docs, stemDocument)
dtm <- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)
head(d, 10)
```

Can big data revolutionize policymaking by governments?

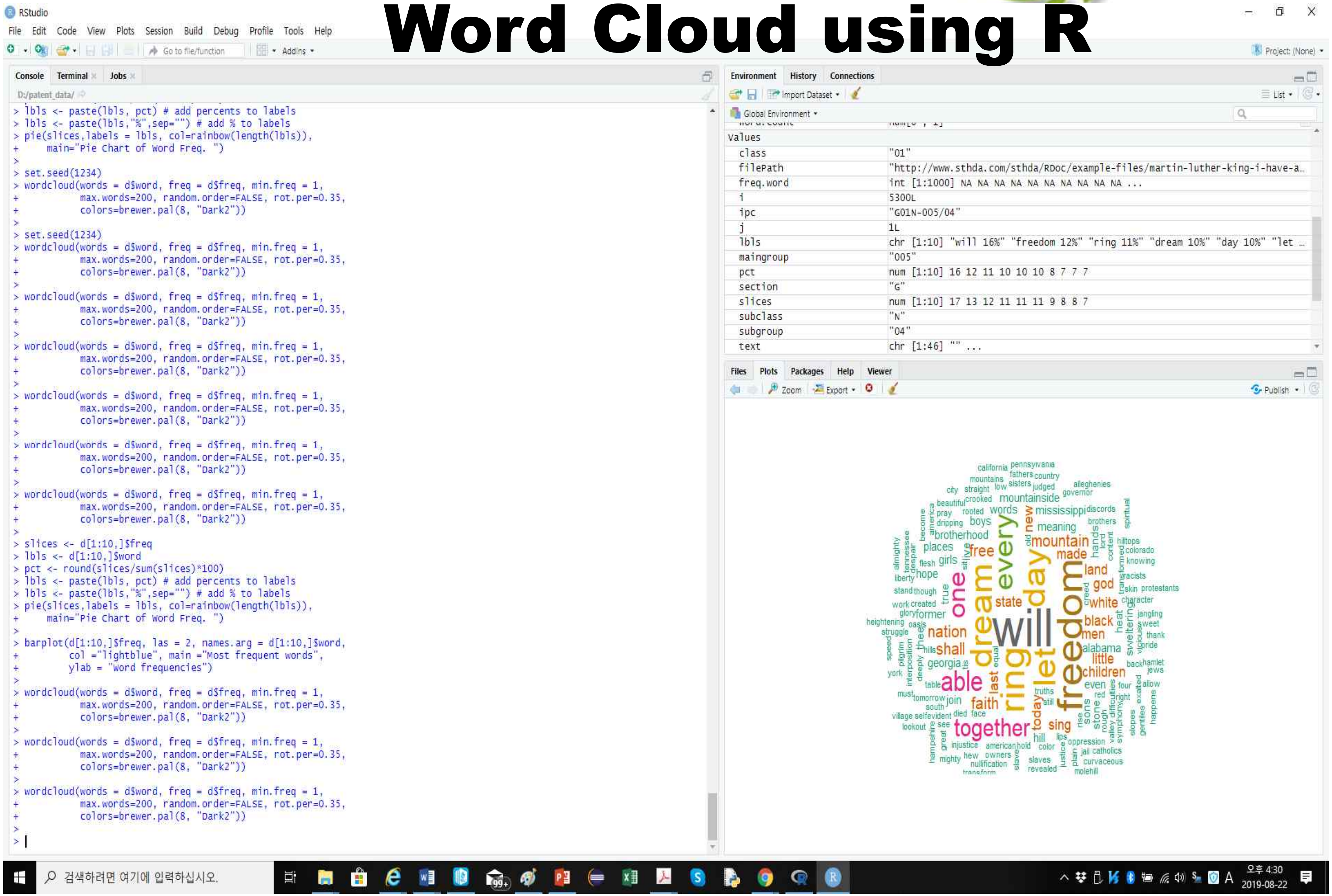
```
set.seed(1234)
wordcloud(words = d$word, freq = d$freq, min.freq = 1,
          max.words=200, random.order=FALSE, rot.per=0.35,
          colors=brewer.pal(8, "Dark2"))

findFreqTerms(dtm, lowfreq = 4)
findAssocs(dtm, terms = "freedom", corlimit = 0.3)
head(d, 10)
barplot(d[1:10,]$freq, las = 2, names.arg = d[1:10,]$word,
        col ="lightblue", main ="Most frequent words",
        ylab = "Word frequencies")
plot(d[1:10,]$freq, d[1:10,]$word, main="Scatterplot Example",
   xlab="Words ", ylab="Freq. ", pch=19)
slices <- d[1:10,]$freq
lbls <- d[1:10,]$word
pie(slices, labels = lbls, main="Pie Chart of Word Freq.")

slices <- d[1:10,]$freq
lbls <- d[1:10,]$word
pct <- round(slices/sum(slices)*100)
lbls <- paste(lbls, pct) # add percents to labels
lbls <- paste(lbls,"%",sep="") # add % to labels
pie(slices,labels = lbls, col=rainbow(length(lbls)),
   main="Pie Chart of Word Freq. ")
```

Can big data revolutionize policymaking by governments?

Word Cloud using R



Can big data revolutionize policymaking by governments?

The strategic difference between dogs and cats for the education-how to solve the problems. Using Big Data?

https://www.youtube.com/watch?v=Oq8nYgnE93Y

- **How about Korea during the big data era?**
  - **Education was a key success factor for the big data era.**

Can big data revolutionize policymaking by governments?

- **So called "Crazy strategy for the business"**
  - **Samsung Electronic Semi Conductor Case**
  - **Hyundai Ship Building Company Case**

Can big data revolutionize policymaking by governments?

Samsung Electronics Semi Conductor Plant located in Korea

Samsung Electronics Announced the Successful Development of 64K D RAM Semi Conductor on Dec. 1, 1983.

KOICA

Can big data revolutionize policymaking by governments?

Samsung Electronics Built Line 3 of Semi Conductor Plant on August, 1987 in Kihung, Korea.

Samsung Electronic Smart Phone Plant located in Vietnam

KOICA

General Yi Sun-Sin and a turtle ship story to get the fund from UK Bank.

Mr. Longbattom, CEO of A&P-Appledore International Ltd (APA), focused on ship design and construction. He let Barclays Bank lend ship building money to Hyundai.

Can big data revolutionize policymaking by governments?

- **In 1983 Hyundai Electronics Industries Co., Ltd. (The name Hynix comes from Hy in Hyundai and nix [from nics in Electronics]) was founded.**
- **Now it became SK Hynix.**
- **Hynix became almost bankrupted, but it makes a lot of profit because of semi-conduct market booming.**

- **Secrets behind Korea`s Economic Success**

https://www.youtube.com/watch?v=bJ0hMr5TSkI

Can big data revolutionize policymaking by governments?

- The world's annual data generation is estimated to be doubling every year, and the overall size will reach 44 zettabytes (that's trillions of gigabytes) by 2020, according to a study by International Data Corporation.
- If all this information was placed in high-end tablet computers, the pile would reach from Earth to the moon more than six times over.

Can big data revolutionize policymaking by governments?

- All this digital data can give you more contemporaneous insight about the economy. The potential is dizzying.
  - Social media feeds can be used to build real-time gauges of sentiment.
  - Satellites in space see which ships dock where and when, whether oil tanks are full or empty, the quality of a crop or even the productivity of a blast furnace.
  - Credit card purchases and email receipts show retail spending.

Can big data revolutionize policymaking by governments?

- Job listings from hundreds of thousands of career sites or corporate websites can reveal employment patterns. And smartphones send location data that show where we are at any given time.
- In time, the “internet of things” could reveal our daily eating habits through web-connected fridges. Satellite images of harbors, car parks, roads and airports can reveal much about the movement of goods.

Can big data revolutionize policymaking by governments?

- Some finance ministries, central bankers and statistics agencies are now starting to dabble in the field in order to understand the economic tides better and more swiftly — a development that could have significant public policy implications. The financial crisis exposed major gaps in official figures.
- The National Bureau of Economic Research's business cycle dating committee, which is the semi-official arbiter of US economic contractions, took until December 2008 — nearly three months after Lehman Brothers went bankrupt — to declare that the US economy had actually entered a recession a year earlier.

Can big data revolutionize policymaking by governments?

- While many economists had concluded as much for some time from the rapidly souring monthly and quarterly data, the statistics did not adequately capture the pace at which the economy was tanking, recalls Diana Farrell, former deputy director of the Obama administration's National Economic Council.
- "The economy was doing a lot worse than we realized, and our policy response was predicated on a much weaker recession," she admits.

Can big data revolutionize policymaking by governments?

- Satellite monitoring of Chinese manufacturing facilities by SpaceKnow.
- How it works: measuring production in China SpaceKnow builds the Chinese Satellite Manufacturing Index from taking millions of snapshots of more than 6,000 industrial sites across China, and uses **artificial intelligence** to turn activity patterns into a numeric measure of how well the country's manufacturing sector is doing.

Can big data revolutionize policymaking by governments?

- Big data could have a “huge” impact on policy, especially around recessions.
- There is a lot that traditional data cannot answer at extreme moments.

II. Big Data and E-government

(Source: Zaher Ali Al-Sai and Laith Mohammad Abualigah, “Big Data and E-government: A review,” 2017 8th International Conference on Information Technology (ICIT))

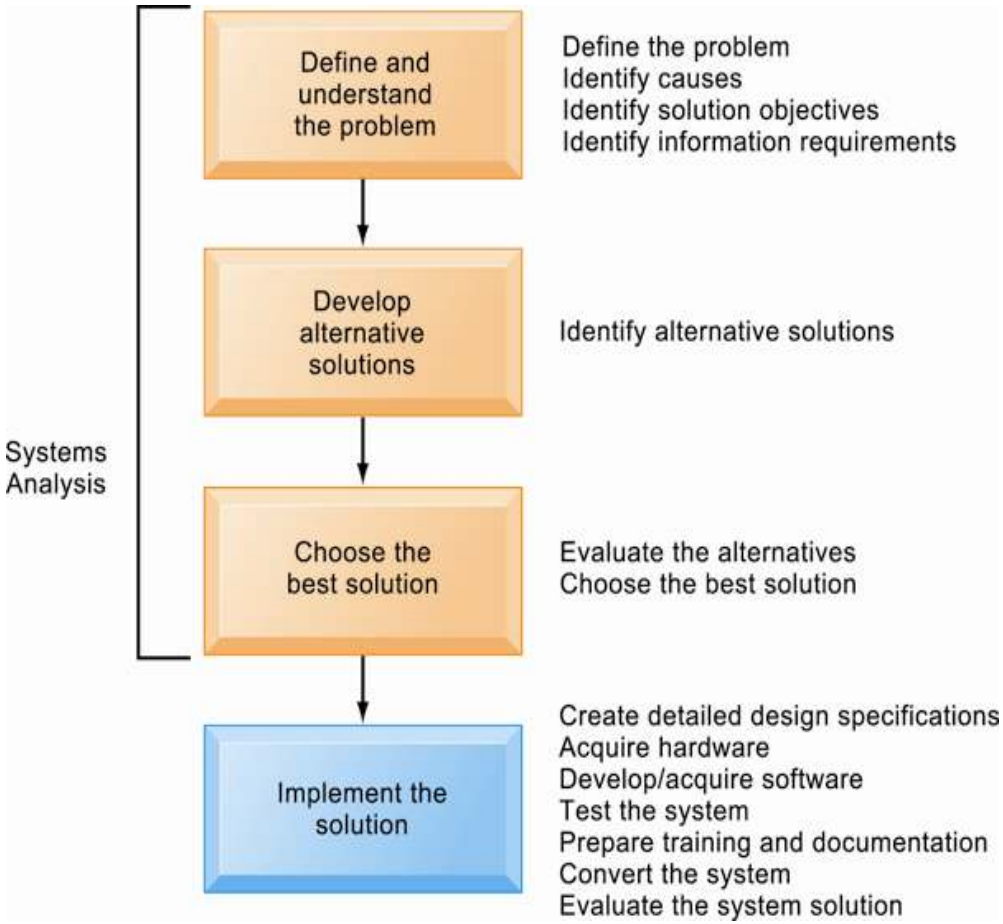
Big Data - the next government affair

Big Data and E-government

- **Big Data for Problem Solving Attitude**

Developing an Information System Solution

Developing an information system solution is based on the problem-solving process.



Problem Solving

- **Earle Dickson** (October 10, 1892—September 21, 1961) was an American inventor best known for creating Band-Aid® brand adhesive bandages.
- He lived in Highland Park, New Jersey, for a large portion of his life.
- Dickson was a cotton buyer at the Johnson & Johnson company.
- His wife, Josephine Knight, often cut herself while doing housework and cooking.

Problem Solving and Systems Development

- Dickson found the gauze stuck to a wound with tape did not stay on her active fingers.
- In 1920, he placed squares of gauze in intervals on a roll of tape, held in place with crinoline.
- James Wood Johnson, his boss, liked the idea, and put it into production.
- In 1924, Johnson & Johnson installed machines to mass-produce the once handmade bandages.
- Following the commercial success of his design, Dickson was promoted to vice president.

Problem Solving and Systems Development

Big Data for Defining and Understanding the Problem

What caused the problem?
Why does it persist?
Why hasn't it been solved?
What are the objectives of a solution?

- **Different people may have different ideas about th e nature of the problem and its severity**
- **Information requirements**
  - Identifies who needs what information, when, where, and how
  - Requirements analysis

Big Data and E-government

- A big challenge for the government of developing national initiatives is to find out how can use the offered opportunities by ICT to promote an effective services. Implementing the
- Big Data in ICT has the power to transform e-government transactions with added value for public services and has the ability to support the digital innovations for e-government

Big Data and E-government

- E-Government can use big data to discover the trends and patterns of peoples' behavior on the social networking so that the government can provide better, effective, efficient service.
- The term of Big Data offers a new opportunities for value creation, discovery , prediction and empower the business intelligence for decision support in e-government.

Big Data and E-government

- Big data can provide a pattern of people's activities and information while the e-government dealing with big data to predict and assume the current needs and satisfaction of their people.

Big Data and E-government

- The researchers can summarize some benefits of having a big data in e-government include the following :
  - Providing and integrating efficient resource of big data.
  - Integrate valued data in e-government to decision making processes .
  - Ability to generate data faster.
  - Higher revenue .
  - Increasing of storage capacity.

Big Data and E-government

- The researchers can summarize some benefits of having a big data in e-government include the following : (Cont'd)
  - Availability of different types data .
  - Empower and Enhance the quality of life .
  - Controlled utilization of E-government resources efficiently.
  - Increasing of transaction processing efficiency.
  - Increase the levels of transparency.

Big Data and E-government

- Challenges include issues related to legal frameworks, policies and principles; data management and protection; identity management and privacy;
- 105 UN Member States have policies and legislation on the right to access e-government information.
- Also the same number produce online policies on open government data and about 113 countries offering online personal data protection Acts and legislation.

III. Strategies for Korean Big Data Applications

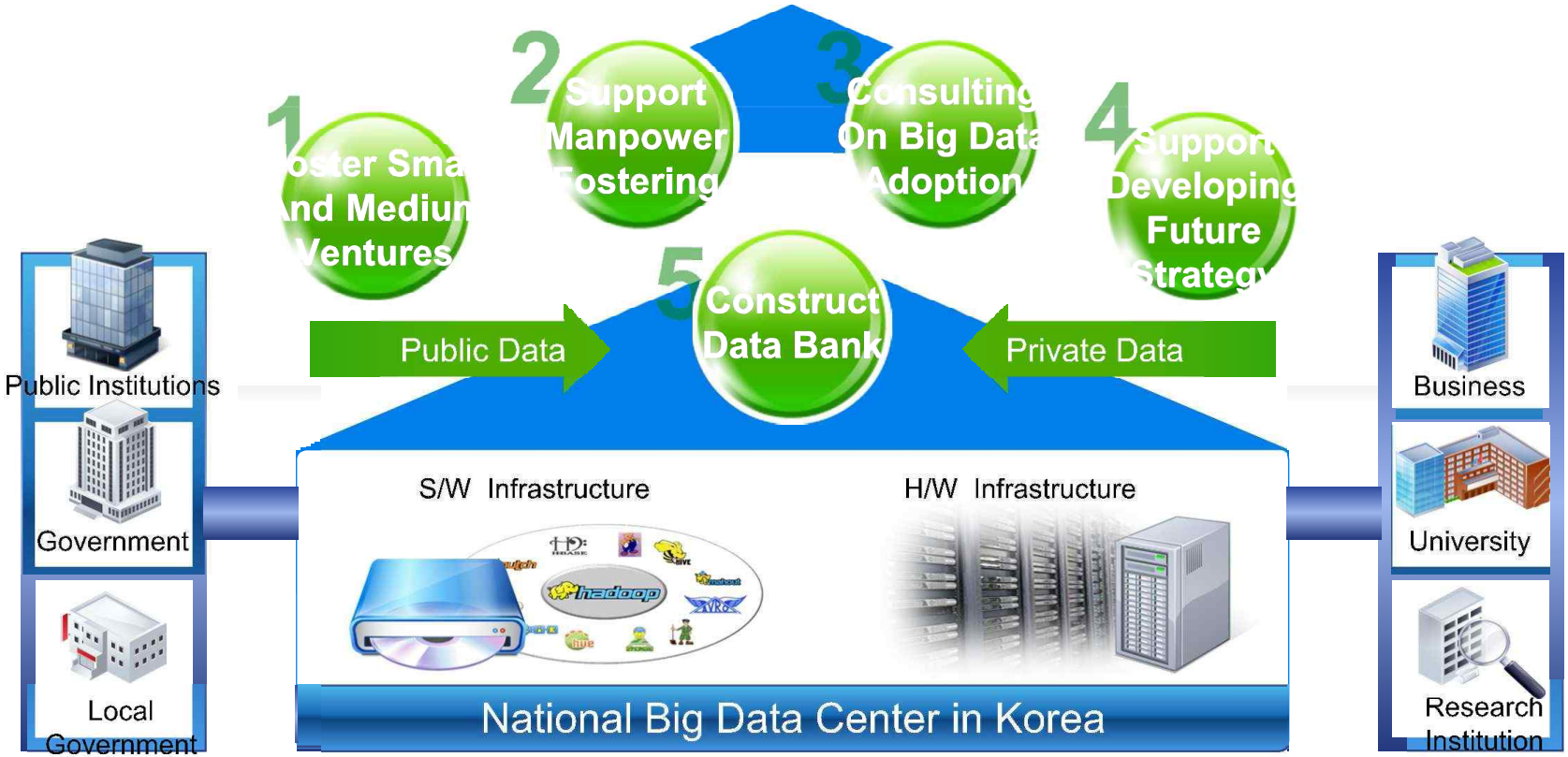
(Source: Kwon Yeong-il (Victor) (Head of National Big Data Center(NBDC)), Big data and its Application in Korea, 2017)

Gov't to inject billions to incorporate smart info tech into state projects.

Strategies for Korean Big Data Applications

National Big Data Center of NIA(National Information Society Agency)

Provide the shared service of big data to the start-up, IT venture companies and academia for fostering bigdata R&D and business



Strategies for Korean Big Data Applications

Supporting Startups and Ventures

Support fostering of data-based startups and
big data technology companies that are small yet strong

Environment for Developers

Provide hardware and software resources that are required in the entire process of service development

Business Idea Competition

Support connecting excellent ideas with startup opportunities

Data Distribution and Exchange Market

Set up the standards and rules for data exchange and develop the market policy for data

Publicizing Excellent Technologies

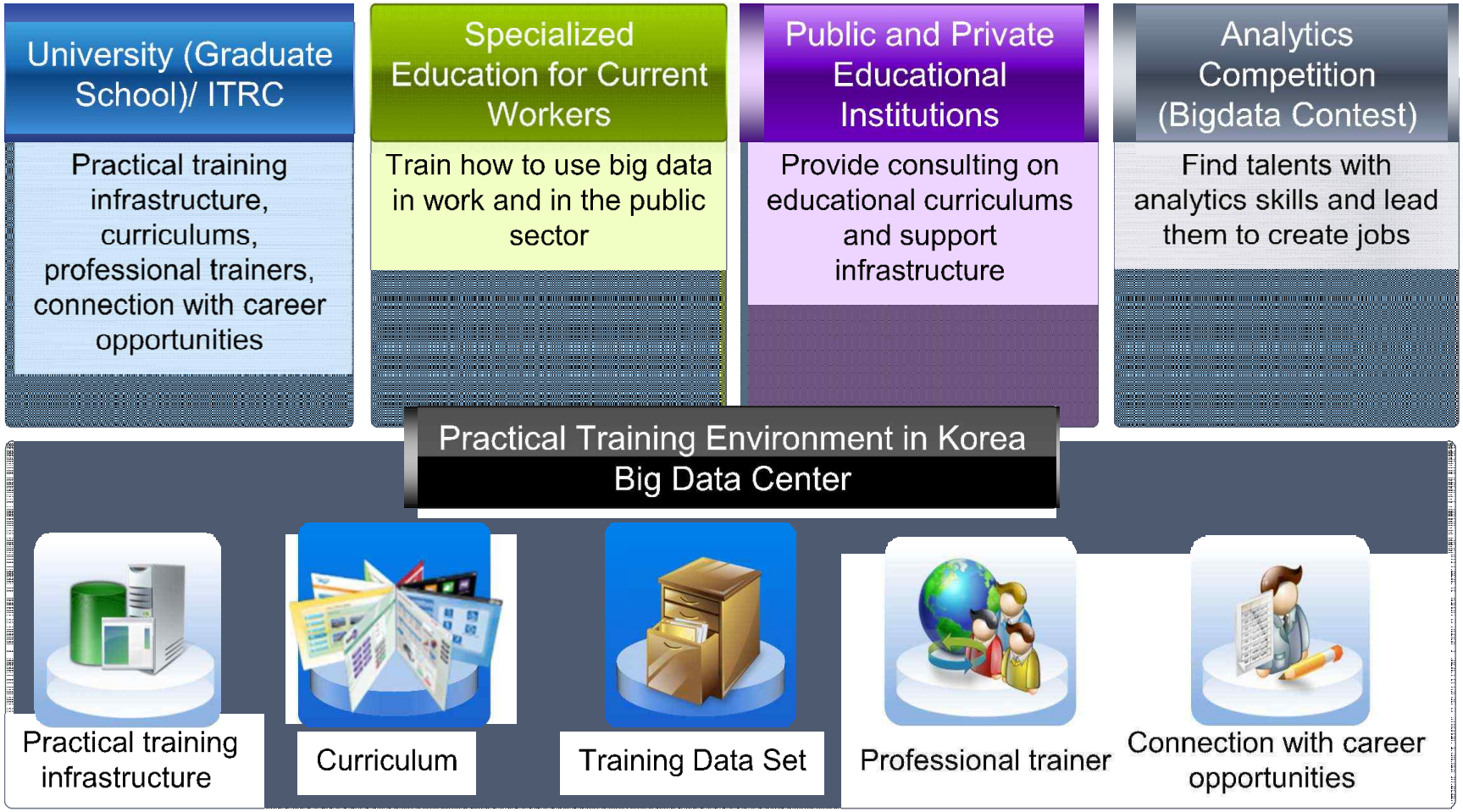
Identify excellent technologies of big data, support their market entry and co-brand the solutions

Test Infrastructure for Big Data Analysis and Use

Provide cluster-based hardware resources considering data access security

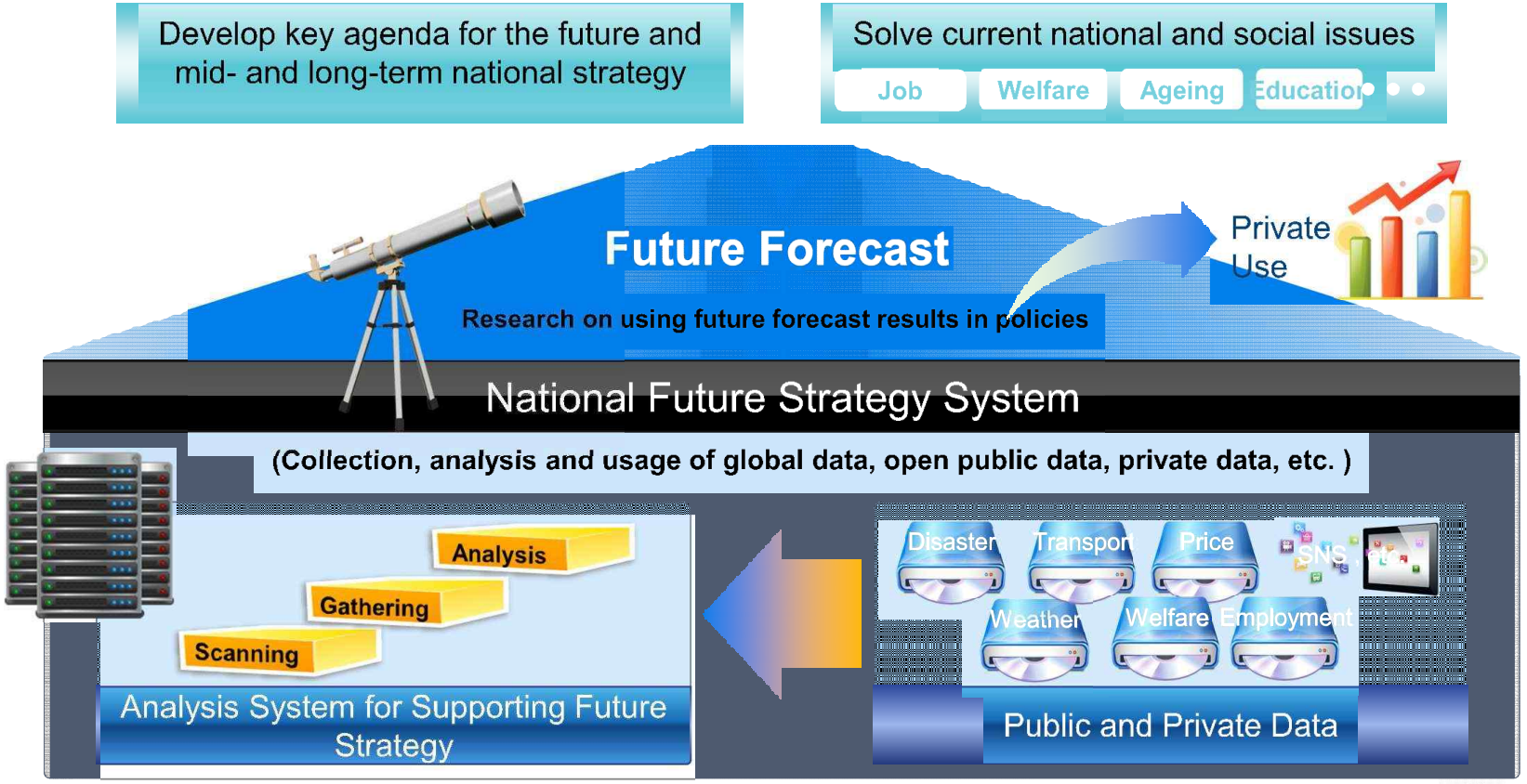
Strategies for Korean Big Data Applications

Fostering Specialized Manpower (Data Scientist)



Supporting National Future Strategy

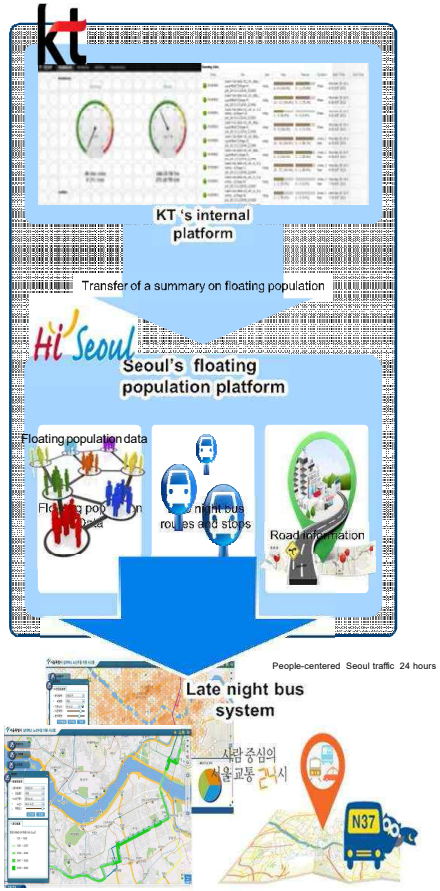
Support the implementation of national future strategy
based on objective evidence including data analytics, knowledge



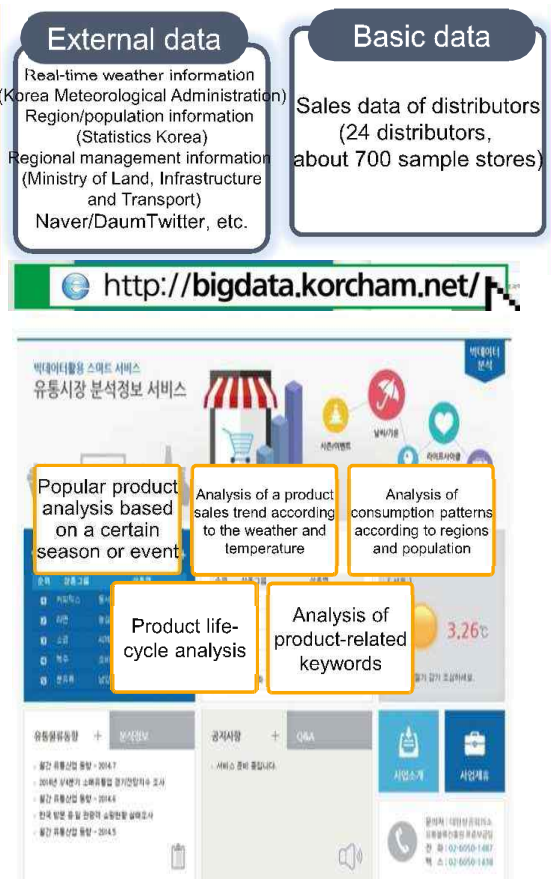
Strategies for Korean Big Data Applications

Case of Big data Flagship Project

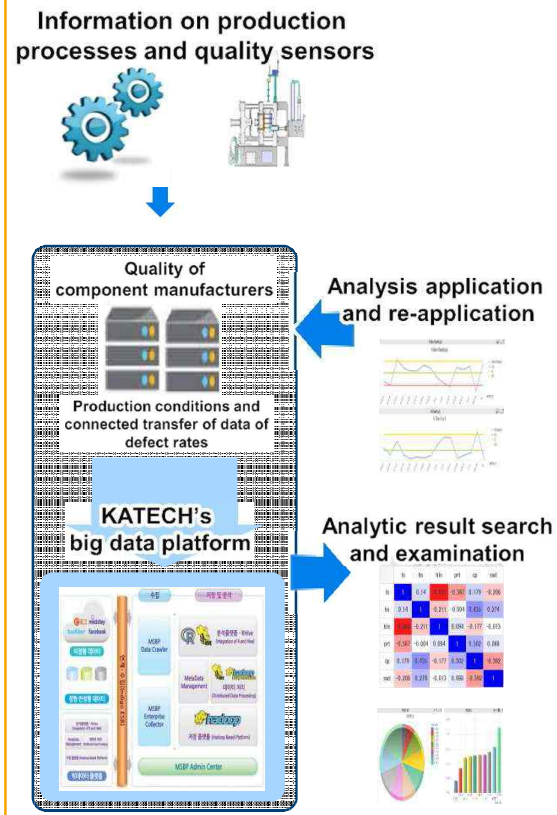
Support for determining the optimal late-night bus route (KT, Seoul)



Distribution market analysis service (Korea Chamber of Commerce and Industry)



Optimization of the manufacturing process of automotive components (Kodaco, Korea Automotive Technology Institute)



Strategies for Korean Big Data Applications

Case 1 : Seoul City Government Case (Night Bus Routing)

In 2013, the City of Seoul conducted a big data analysis on telecom data, discovering that people called a lot looking for a ride after midnight. The City of Seoul launched Night Bus Line after midnight.

Economic Impact

Benefits to bus transportation companies and the citizens in Seoul

Economic impact of hundreds of million (won) through the night bus service

Application and Follow-up

- The best bus route based on the data
- Propagation to other cities including Busan, Kwangju, and so on.

Case 1 : Seoul City Government Case (Night Bus Routing)

System Architecture

KT + Seoul City Government

kt

KT Platform

Hi Seoul

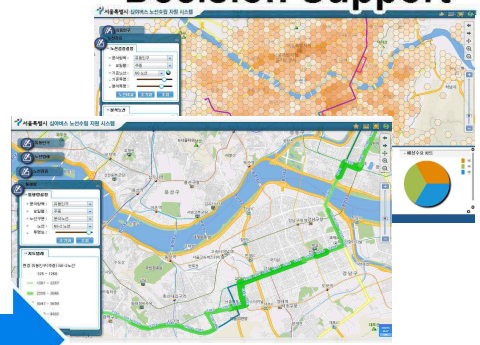
Seoul Pedestrian Flow

PEDERSTRIAN ESTIMATION

NIGHT BUS LINE, STATION

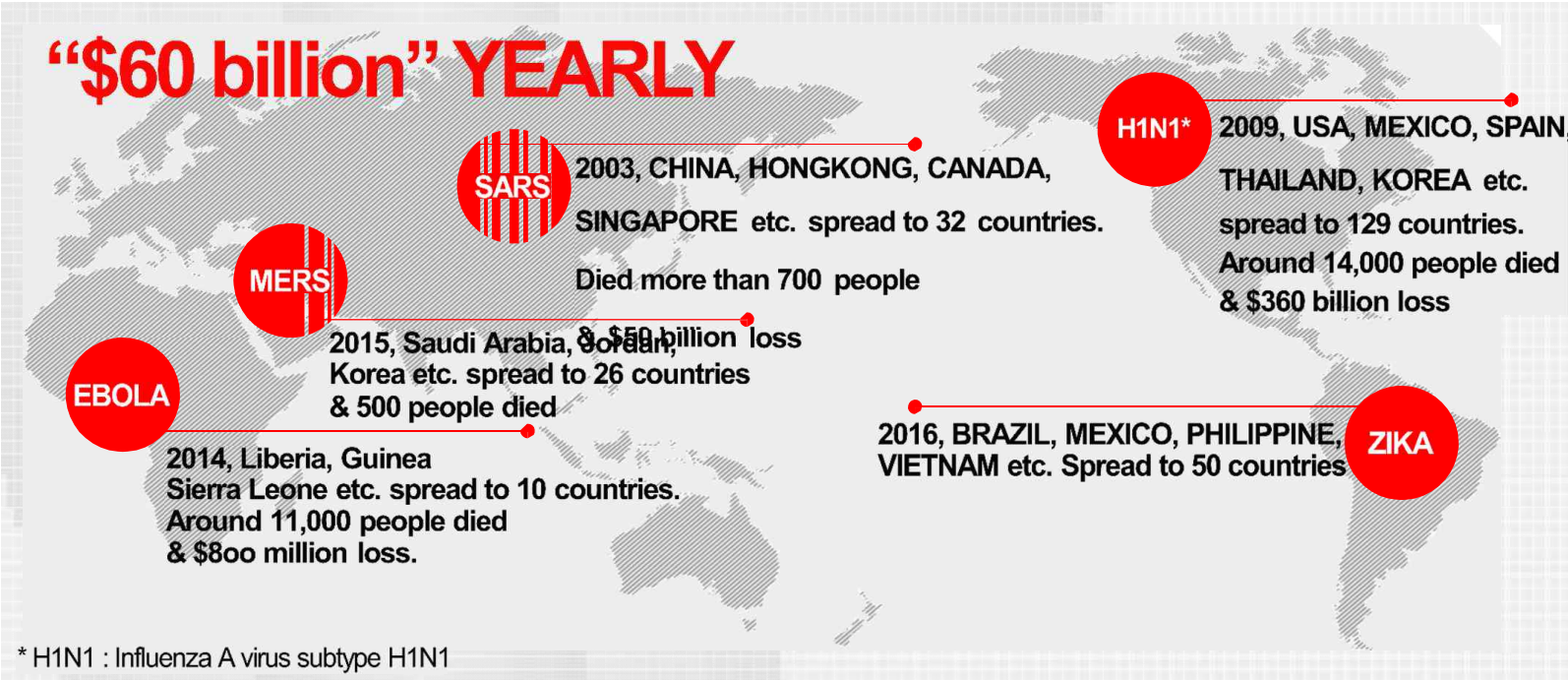
ROAD

Bus line Routing Decision Support



NIA NATIONAL INFORMATION SOCIETY AGENCY

Case 2 : Infectious Diseases Monitoring

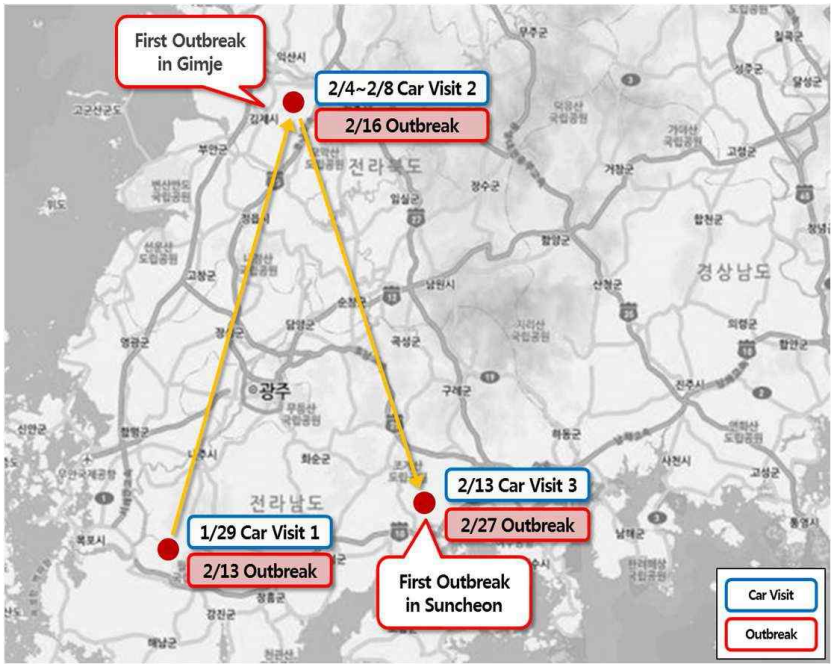


" **Within 100 years, at least more than one of the infectious disease around the world expected to outbr eak more than 20% of possibility and as a result, economical loss will be over $60 billion yearly.**

-SOURCE : Global Health Risk Framework Report (Jan. 2016)

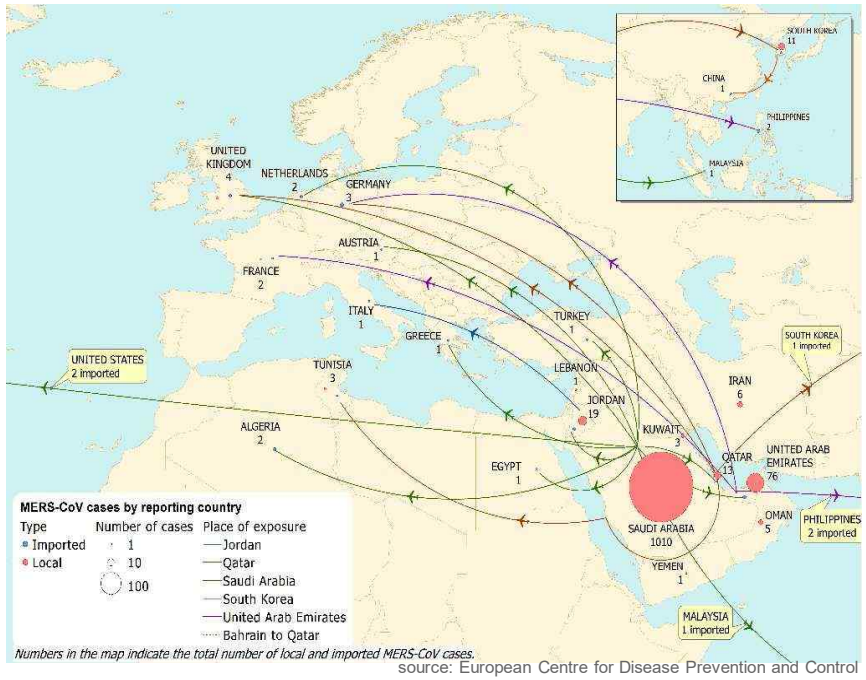
Case 2 : Infectious Diseases Monitoring

Cause of Animal Epidemic Transmission



Spread by vehicles visiting the infected farm

Cause of Human Epidemic Transmission



source: European Centre for Disease Prevention and Control

Spread by people visiting the infected area

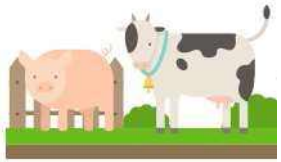
Using big data and ICT technologies can help prevent infectious disease from spreading

Case 2 : Big data flagship pilot project

3. Korea's Case : Animal Epidemic

Data Source

Movement log of livestock vehicles collected via GPS



Government information on Farms(e.g. Livestock information, farm visiting history)

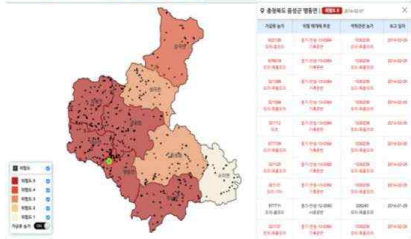
Predictive Analytics

농림축산검역본부 Animal and Plant Quarantine Agency

Predict modeling where animal diseases will occur with Deep Learning

Utilizing Information

KAHIS 국가동물방역통합시스템



Provide predictive information on where the epidemic animal diseases will occur

•https://www.youtube.com/watch?v=KoZK_0vF3Us&feature=player_embedded

Case 2 : Infectious Diseases Monitoring

4. History of Preventing Infectious Disease

Ministry of Science, ICT and Future Planning

Ministry of Agriculture, Food and Rural Affairs

Ministry of Health and Welfare

with kt

	Animals Epidemic	Human Epidemic
2014	Create a model to predict the infectious disease transmission utilizing data getting from the vehicles that carry live stocks	
2015	Build a prediction system in 'Animal and Plant Quarantine Agency' and leverage predictive models to prevent the spread of AI(Avian Influenza) & FMD(Food-and-mouth disease)	Deliver consultation to block inflow of infectious diseases using roaming data
2016		Develop 'Smart Quarantine System' to avoid infectious disease using roaming data from KT
2017	Plan to add farm entry log in order to upgrade the predictive model	Schedule to add the roaming data from SKT and LGU+ to 'Smart Quarantine System'

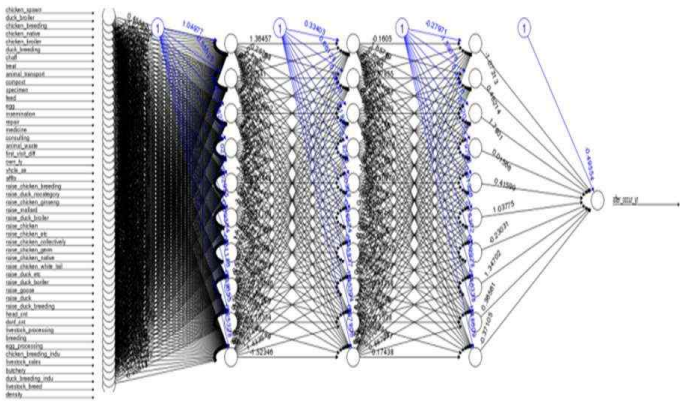
* KT, SKT, LGU+ are telecommunication companies from South Korea

Case 2 : Infectious Diseases Monitoring

5. Impact of Korea's Case (Animal Epidemic)

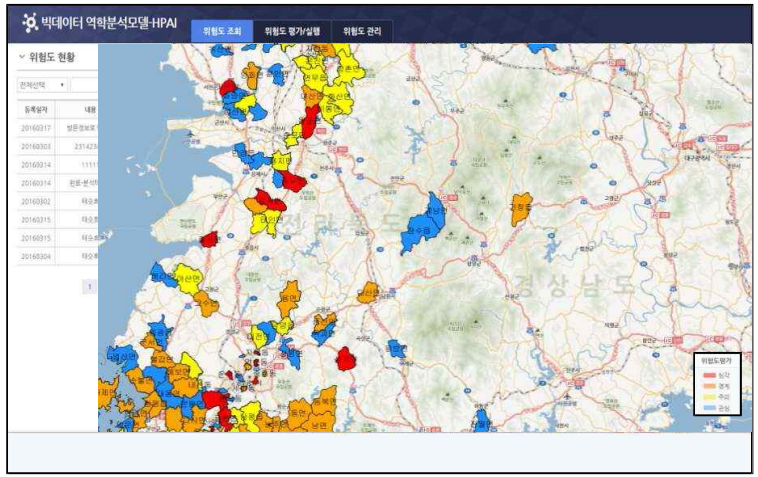
The predictive model has been used every time an infectious disease occurs

- The model predicted 97 AI(Avian Influenza) cases out of 116 cases (83.6%) from the 2nd half of 2015 to the 1st half of 2016
- It helped cutting $ 76 Million losses
- This model also has been applied to other infectious disease cases such as FMD(Foot and Mouth Disease)



Predictive Model
(developed by machine learning)

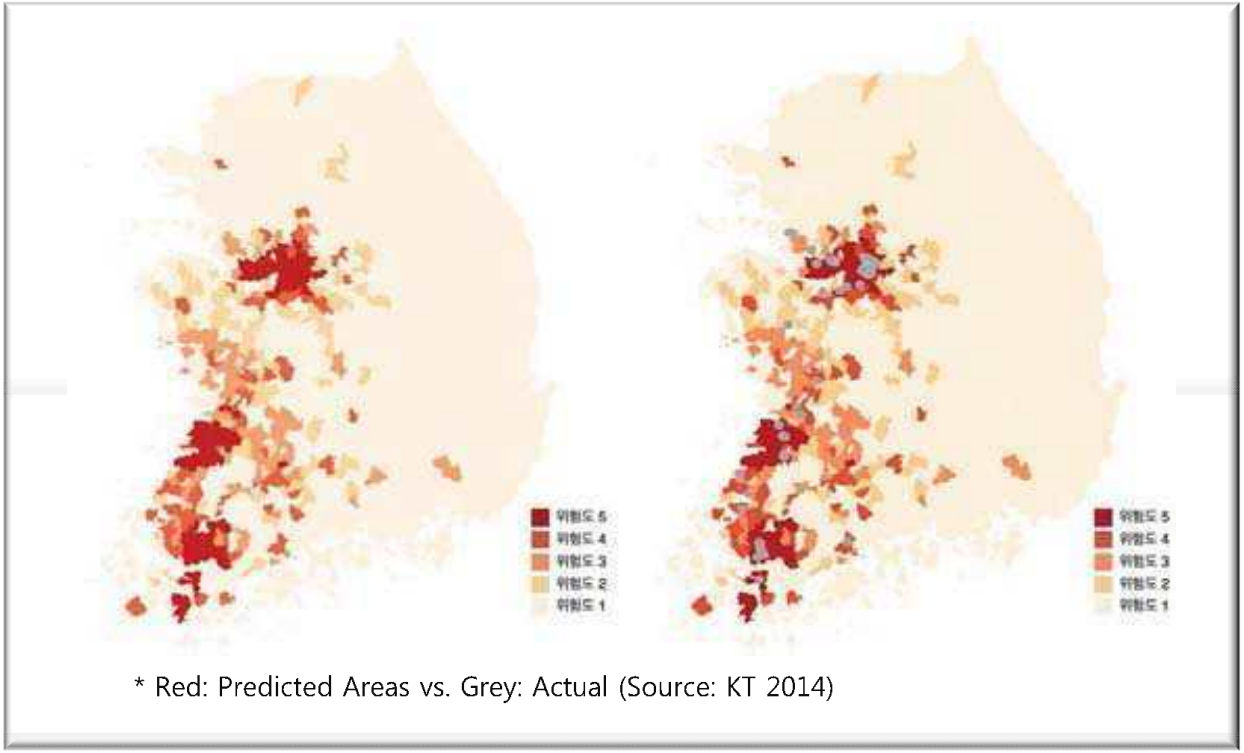
Result Display & Use



Case 2 : Infectious Diseases Monitoring

In 2014, KT conducted a big data analysis with KAHIS to predict the spread of AI (Avian Influenza) in Korea.

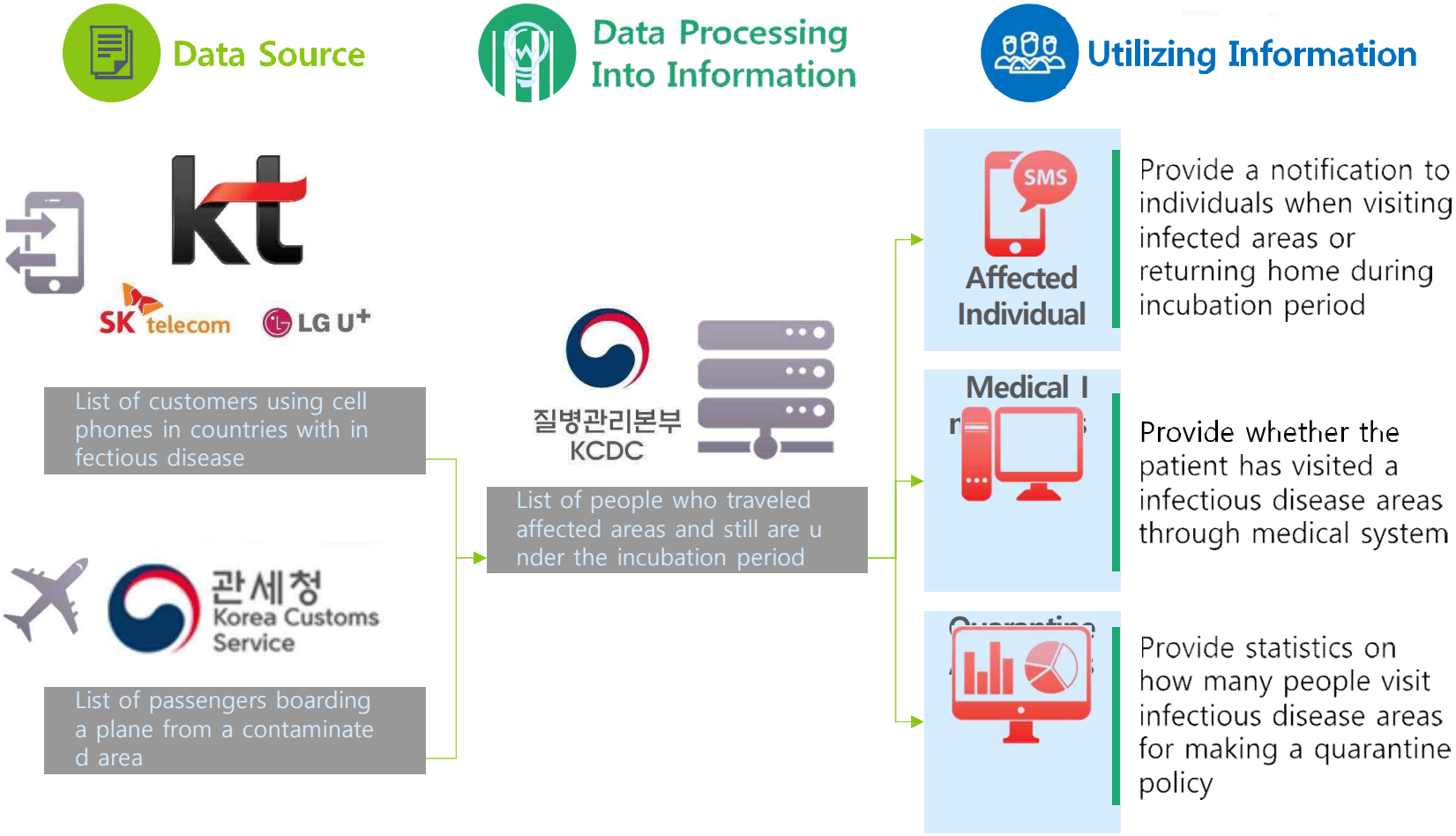
Prediction and actual breakout regions



* Red: Predicted Areas vs. Grey: Actual (Source: KT 2014)

Case 2 : Infectious Diseases Monitoring

6. Korea's Case: Human Epidemic

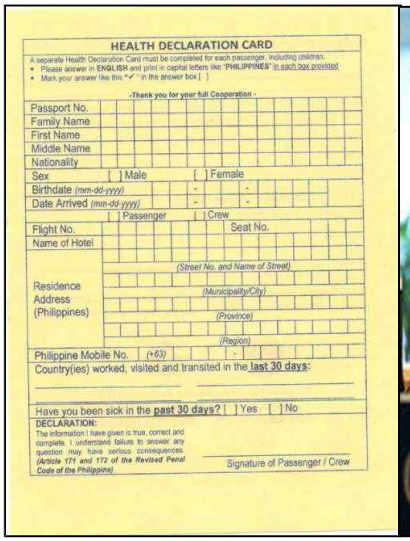


Case 2 : Infectious Diseases Monitoring

6. Korea's Case: Human Epidemic

KT and KCDC have started 'Smart Quarantine Service' since November 17th , 2016

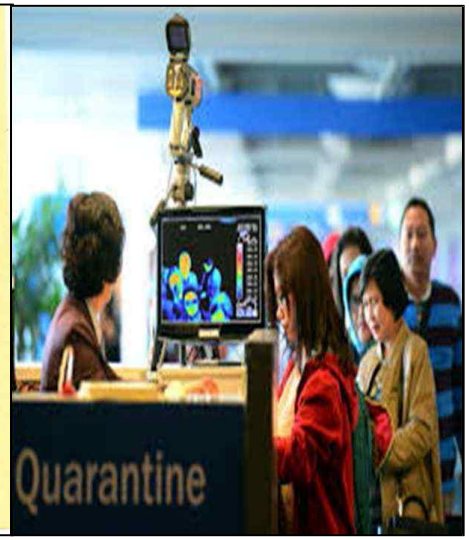
- **Send 630K+ individuals a warning message of Infectious Diseases**
- **Find 36K+ individuals who stop over a clean country after visiting affected areas** (It is difficult for the quarantine agency to find this case)
- **SKT and LGU+ subscribers will receive a service from April 2017**


HEALTH DECLARATION CARD

A separate Health Declaration Card must be completed for each passenger, including children.
- Please answer in ENGLISH and print in capital letters like "PHILIPPINES" in each box provided
- Mark your answer like this "✓" in the answer box [ ]

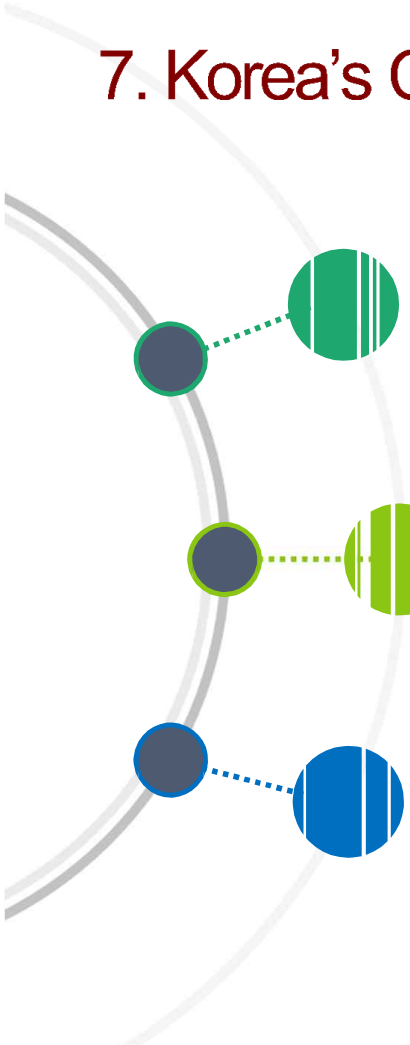
-Thank you for your full Cooperation -

Passport No.
Family Name
First Name
Middle Name
Nationality
Sex [ ] Male [ ] Female
Birthdate (mm-dd-yyyy)
Date Arrived (mm-dd-yyyy)
[ ] Passenger [ ] Crew
Flight No. Seat No.
Name of Hotel
(Street No. and Name of Street)
Residence Address (Philippines)
(Municipality/City)
(Province)
(Region)
Philippine Mobile No. (+63)
Country(ies) worked, visited and transited in the last 30 days:
Have you been sick in the past 30 days? [ ] Yes [ ] No
DECLARATION:
The information I have given is true, correct and complete. I understand failure to answer any question may have serious consequences. (Article 171 and 172 of the Revised Penal Code of the Philippines)
Signature of Passenger / Crew



Case 2 : Infectious Diseases Monitoring

7. Korea's Case: Human Epidemic



Introduce Big Data Practices in Public Administration

- Introduce the best case to prevent the infectious diseases from spreading by adding big data technology to the existing quarantine system

Share ICT Infrastructure and Big Data Utilization Method with Vietnam

- Share the way of constructing ICT infrastructure to prevent infectious disease from spreading
- Share the way of using big data and developing big data model to improve Korea's quarantine system (both human epidemic and animal epidemic)

Provide Consultation and Training ICT Infrastructure Service

- Provide consulting and training services for the big data model and systems to prevent national animal epidemic from spreading
- Provide consulting and training services for visitor management system to prevent infectious disease from spreading

Case 3 : Target Business Analysis

In 2015, BC Credit Card conducted an analysis on structured and unstructured data in combination, producing better marketing campaign outcome.

Social Data

Rank	Keyword	문서수	빈도수
1	조카	17,880	24,901
2	할머니	16,671	61,561
3	엄마	13,666	62,566
4	어머님	12,200	18,404
5	할아버지	7,855	11,947
6	아빠	7,326	29,099
7	아버님	6,358	9,659
8	가족	5,612	10,455
9	언니	5,129	19,613
10	삼촌	4,592	6,441
11	신랑	3,712	7,677
12	동생	3,530	5,910
13	오빠	3,312	7,506
14	부모님	3,212	4,817
15	남편	3,043	6,334
16	고모	2,976	4,368
17	어머니	2,731	5,655
18	부모	1,874	2,750
19	아버지	1,870	4,062
20	아내	1,256	3,198

Convergence

Rule Based Mapping

Demography

Key

Industry/Merchants

Purchase Time

Purchase Volume

Job

Purchase Reason

Card Data

가맹점니즈 업종코드	가맹점니즈업종 _대분류명	가맹점니즈업종 _중분류명	가맹점융합 DB업종명
N033	사교육		서적
N1022	생활쇼핑	편의점/슈퍼마켓	문구용품
N034	서적문구		문구용품
N042	스포츠		자전거
N122	여행		워터파크
N122	여행		놀이공원
N0512	영화공연	공연	박물관
N043	오락		실내놀이터
N131	유흥		피자
N1411	육아출산	유아교육	어린이집
N1412	육아출산	유아용품	완구
N1412	육아출산	유아용품	유아용품
N1412	육아출산	유아용품	아동복
N1412	육아출산	유아용품	아동의류
N1412	육아출산	유아용품	완구점
N1412	육아출산	유아용품	캐주얼의류
N1412	육아출산	유아용품	장난감
N112	음료제과		피자
N1122	음료제과	커피	피자
N1122	음료제과	커피	햄버거
...	...	...	...

Target Business Derived

Strategies for Korean Big Data Applications

Media Broadcasting on Big Data Project

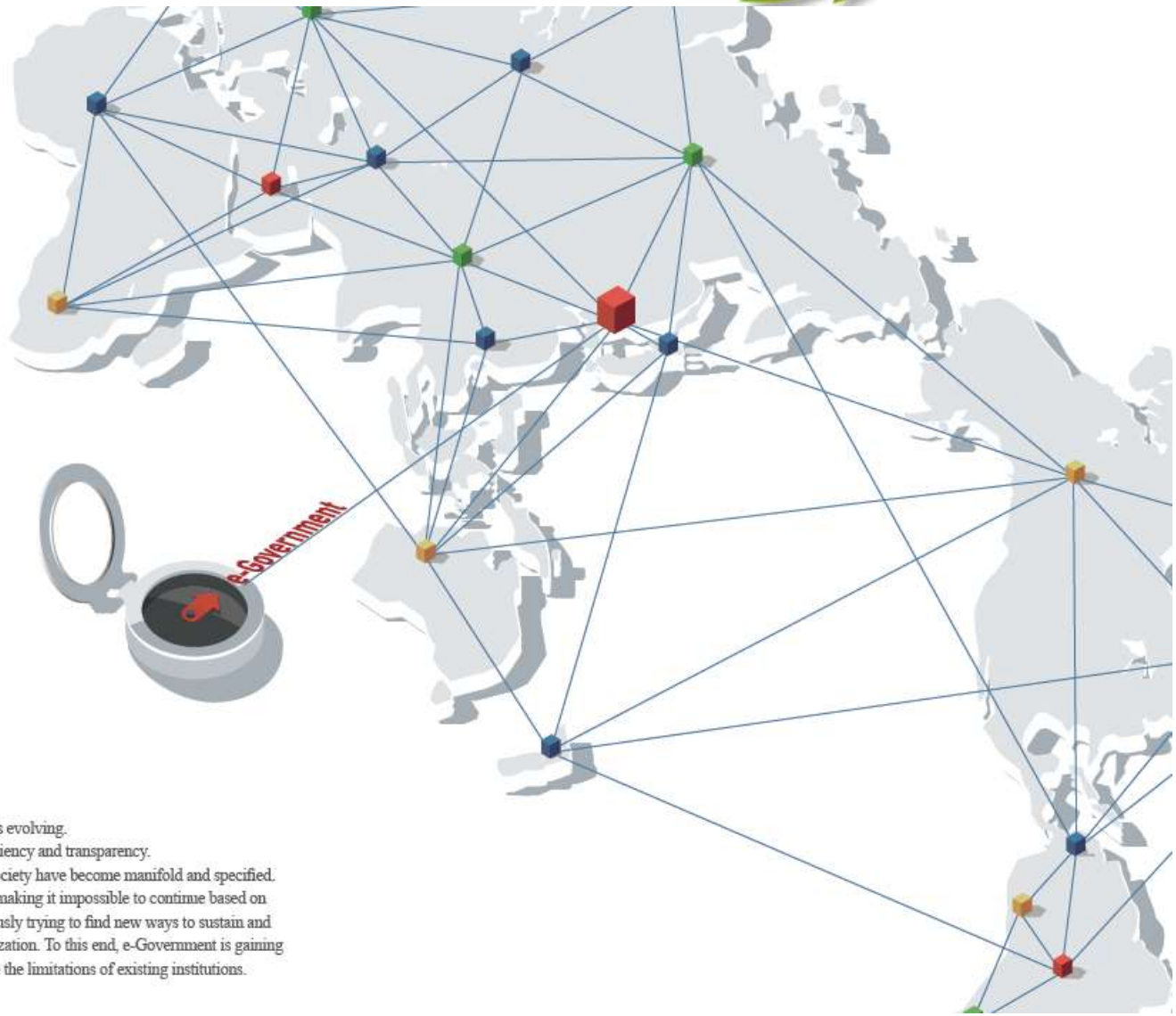
Seoul Bus Routing (2013)

Big Data Consulting Projects (2014)

IV. Big Data and Artificial Intelligence for e-Government

(Source: 1. Korean Government, Leading_the_world_e-Government (brochure)
2.The presentation by Dr. Jim Spohrer (IBM) © IBM UPWard 2016)

Big Data and Artificial Intelligence for e-Government

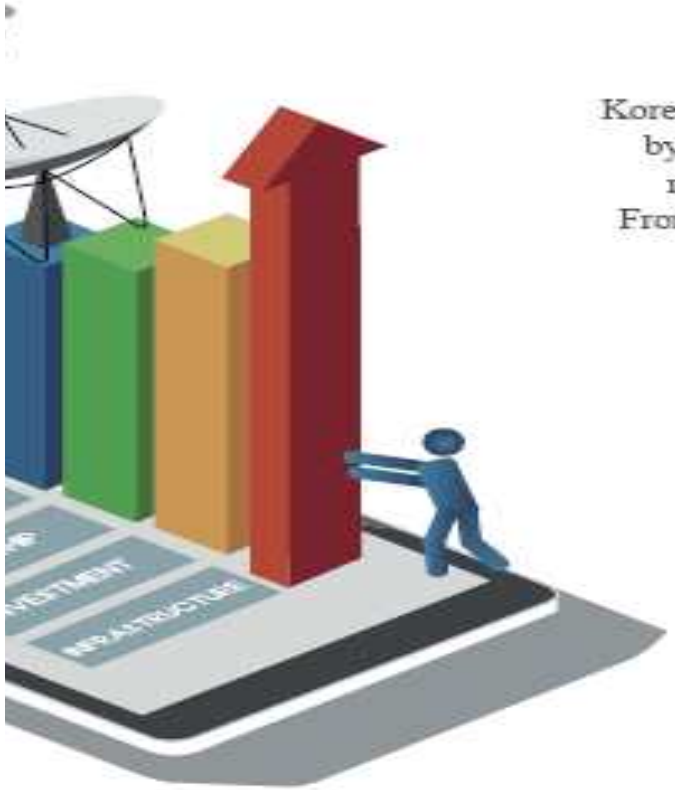


Global order created by informatization, and the evolving role of government

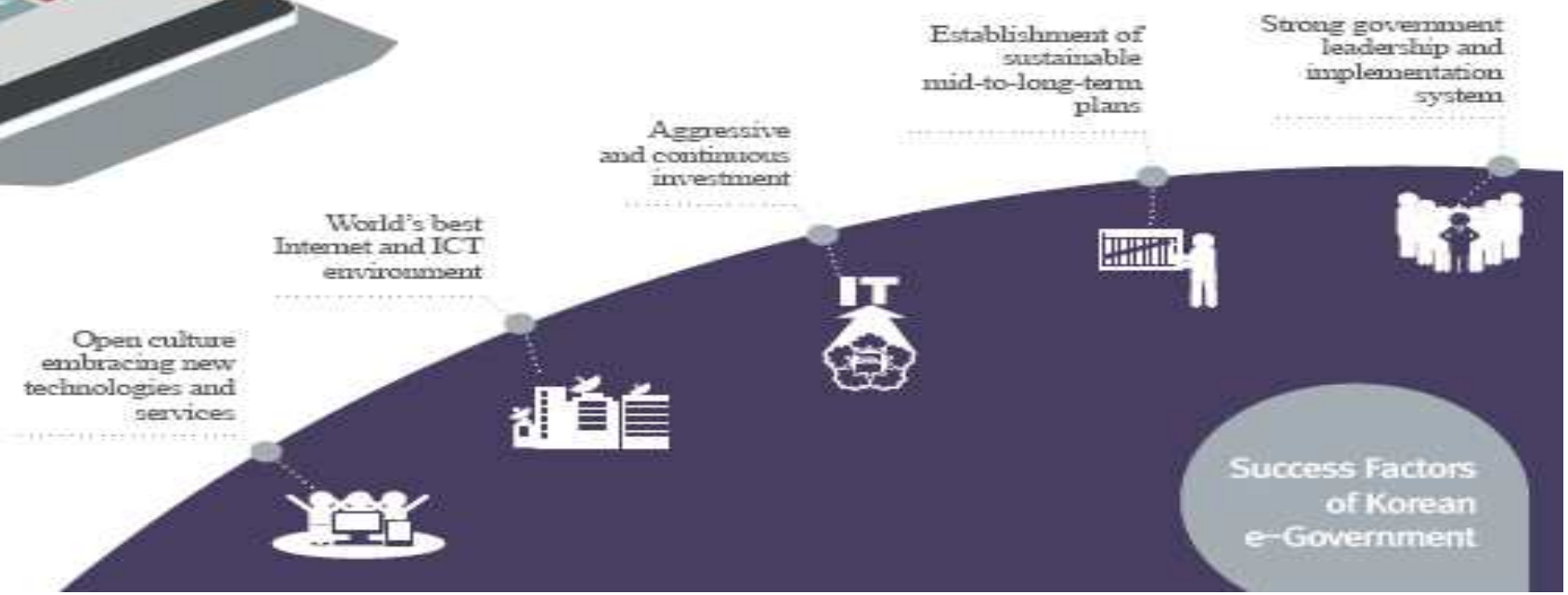
As democracy and ICT develop, the role of government is evolving.
People are expecting a higher level of administrative efficiency and transparency.
The policy requests made by a diverse spectrum across society have become manifold and specified.
Now, people are asking for a more evolved government, making it impossible to continue based on the existing method. The Korean government is continuously trying to find new ways to sustain and develop amidst the new global order created by informatization. To this end, e-Government is gaining attention as a new form of government that can overcome the limitations of existing institutions.

Big Data and Artificial Intelligence for e-Government

This is why Korean e-Government is special.



Countries around the world are promoting e-Government projects by applying ICT in their administrations, and are achieving great results. Korea is also one of the countries that has been able to achieve national development by establishing e-Government. However, performance and results are not the only reasons why the international society is focusing on the e-Government of Korea. From the ruins of war in the 1950s, Korea achieved the "Miracle of the Han River," through remarkable economic development in only half a century. Not only that, Korea became the first country to turn itself from a recipient to a donor country. For this, Korea made a bold decision to invest 1% of its national budget in e-Government. Through this, Korea was able to realize a convenient, transparent, and efficient government, faster than any other country around the world. e-Government was the driving force behind Korea's rapid development.



Big Data and Artificial Intelligence for e-Government

Pioneering the most solid and fastest path to e-Government

Korea built its national basic computer network and high-speed ICT infrastructure before IT was proliferated. IT was the key to the foundation of sustainable development. This foundation became the driving force in overcoming the 1997 Asian financial crisis. Back then, e-Government was the way to fundamentally innovate the overall governance restructuring and operation system of the country.
Now, e-Government is applied not only to administrative tasks, such as HR, finance, and procurement, but also to overall lifestyle, including corporate support and civil services. It also includes a system where people can freely take part in the policy-making process, realizing 'e-democracy' to bring about a nation run by the people.
Korea's e-Government was developed based on this solid foundation and is recognized as an unprecedented global success case.

-1997

The establishment of G2G-based e-Government for efficient administrative work

- Began issuing resident registration numbers, a key factor in e-Government, by amending the Resident Registration Act (1968)
- Government-led development and distribution of Korean-made computers (1970s)
- Launched the first government organization related to IT (April 1970)
- Promoted the computerization of administrative works, such as procurement, communications, customs, security, civil affairs, judicial precedents, etc (1970s)
- Established the national computer network and connected it to the local government systems (1990s)

1997-2012

The expansion of convenient, G4C-based e-Government

- Enacted the "e-Government Act" and established the "e-Government Special Committee" (2001)
- Established the security system for e-authentication (2000)
- Expanded the foundation for e-Government services high in demand, such as national tax (2001), educational administration (2002), food and drugs (2003), and employment (2004)
- Promoted easily accessible m-Government (2011)
- Enacted Personal Information Protection Act in consideration of the adverse effects of e-Government

2012-2016

e-Government that connects and converges departments and services

- Ranked 1st place in three UN E-government Surveys (2010, 2012, 2014)
- Enhanced customized one-stop services for childbirth, inheritance, welfare, and year-end tax adjustment (2010)
- Improved the transparency of administrative services by disclosing public data and original administrative documents
- Advanced e-Government services through next-generation technologies (IoT, cloud computing, big data, and mobile)
- Established Korean e-Government Master Plan 2020 (2016)

KOICA

숭실대학교

Big Data and Artificial Intelligence for e-Government

With

e-Government 2020,

Korea will take a leap into the future.

The e-Government environment is rapidly changing. New ICT will create a hyper-connected society that connects people, things, and space without any constraints. In particular, new technology converged with IoT, cloud, big data, and mobile, the so-called 'ICBM', is transforming the future of e-Government.

The blueprint for the new Korean e-Government that will be realized within the next 5 years is an e-Government that is sensitive to the experiences of citizens, providing services meticulously customized to individuals; an e-Government that realizes advanced intelligence administration based on new technologies and builds an intelligent pan-governmental collaborative environment; and an e-Government that creates new sources of wealth by sustainable digital 'new deals' that foster the new digital industry.

Korea's new e-Government 2020 is a human-focused e-Government realized by advanced ICT.

Ranked 1st place in three consecutive UN E-government Surveys
The best e-Government recognized by the world

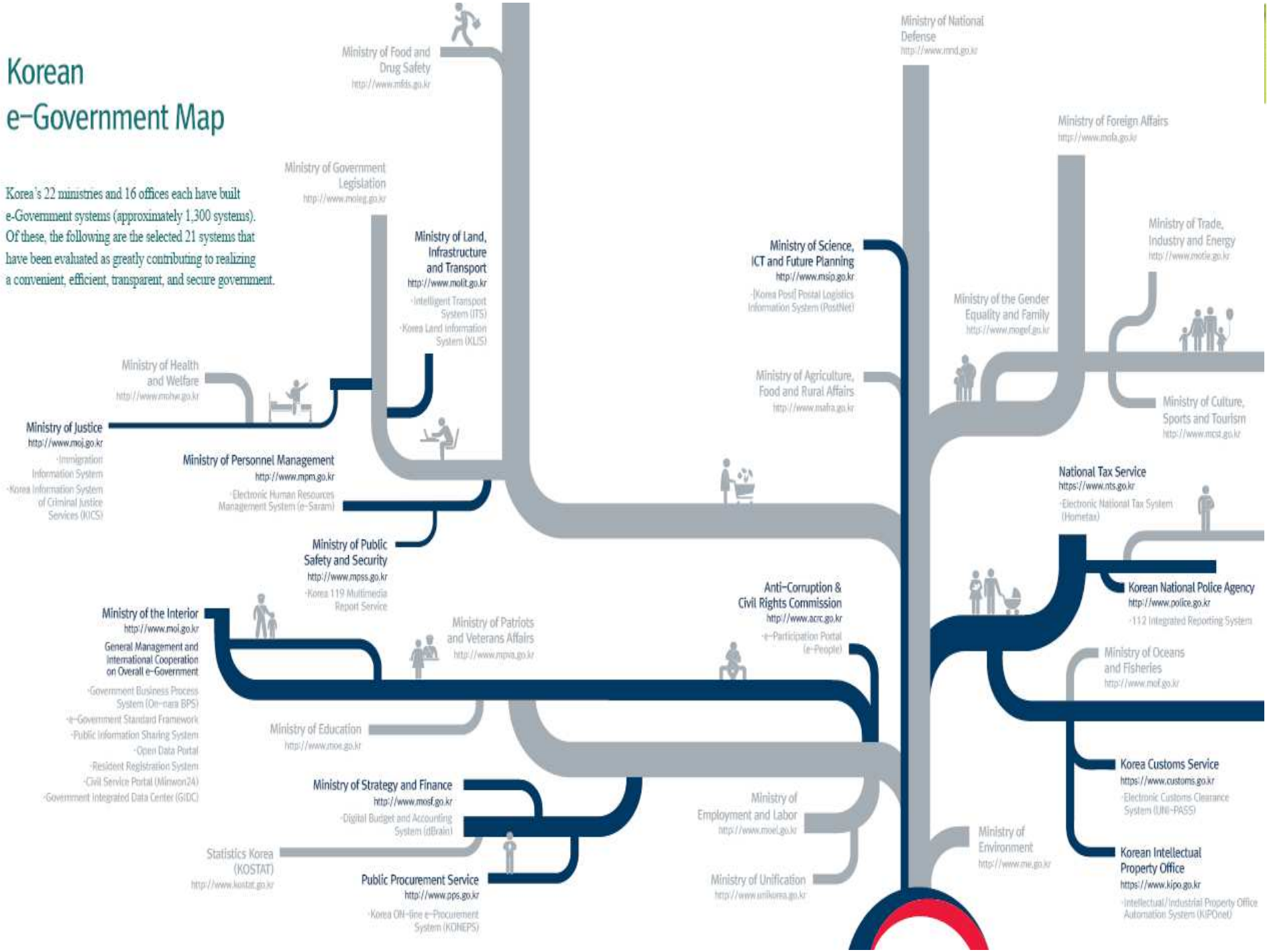
Korean e-Government

The world's best e-Government of Korea has ranked 1st place in three consecutive UN E-government Surveys from 2010. This is an unprecedented record in the history of UN E-government Surveys. Also, the various systems built by Korean e-Government are exported to many countries, receiving excellent feedback. The systems of the Korean e-Government are recognized for their excellence, receiving many awards from diverse international organizations, including the prestigious UN Public Service Awards, widely regarded as the Nobel prize of public service.

Year		2010	2012	2014
Overall		1st	1st	1st
	Online service index	1.00(1st)	1.00(1st)	0.97(3rd)
	Information communication infrastructure index	0.64(13th)	0.83(7th)	0.93(2nd)
	Human capital index		0.94(6th)	0.92(6th)

Korean e-Government Map

Korea's 22 ministries and 16 offices each have built e-Government systems (approximately 1,300 systems). Of these, the following are the selected 21 systems that have been evaluated as greatly contributing to realizing a convenient, efficient, transparent, and secure government.



- **Convenient e-Government**
  - **Realizing e-Government that makes people's lives convenient e-Gov.**
  - **Korean e-Government provides online administrative services at anytime, anywhere in order to make people's lives more convenient.**
  - **All Koreans can use administrative civil services whenever and wherever they want, and can notify, report, and pay a variety of taxes online.**
  - **Also, quality of life is improved with the application of postal logistics information, traffic information, and patent information provided by e-Government.**

- **Transparent e-Government**
  - **Korean e-Government is realizing transparent administration and increasing the people's right to know by disclosing the administrative process of tasks online that are closely related to people's lives.**
  - **The public procurement and customs clearance processes are computerized online to ensure fairness and transparency.**
  - **The criminal justice process is also disclosed online to gain people's trust.**
  - **Also, people's voices are actively reflected in policy formulation.**
  - **Public information is disclosed to the people to help drive new growth engines.**

- **Efficient e-Government**
  - **Korean e-Government is realizing effective administration.**
  - **Efficiency of administrative tasks are improved by building systems such as On-nara, where all administrative tasks are done online, dBrain, where national finances are comprehensively managed real-time, and e-Saram, enabling performance based HRM.**
  - **Also, by providing a standardized framework for software development, the burden of overlapping development was diminished, and administrative task processing time was dramatically shortened due to sharing administrative information.**
  - **Through this, costs, human capital, and time were saved, and work efficiency was improved.**

- **Secure e-Gov.**
  - **The Resident Registration System that provides 24/7 service, and Government IDC, the world's first governmental data center, are the backbone to seamless government tasks even in the event of unexpected disasters.**
  - **In terms of cross-border security, the Immigration Information System blocks entry by dangerous individuals preventing potential terrorist incidents and strengthening national safety.**
  - **Internally, 112 for emergency and 119 for disaster, are keeping all Koreans safe.**

- **Korea will share its knowledge and experience of e-Government and contribute to the development of the global e-Government.**

Big Data and Artificial Intelligence for e-Government

Korea will share its knowledge and experience of e-Government and contribute to the development of the global e-Government.

Korea is sharing its e-Government experience with the rest of the world to contribute to bridging the digital divide between countries. Requests for cooperation and exchange are continuously coming from not only developing countries, but also major international organizations such as WB, ADB, IDB and so forth. From 1998 to 2014, a total of 1,400 people from 111 countries learned about Korea's e-Government through local training and invitational workshops.
The e-Government of Korea will share its accumulated experiences and knowledge with the world to solve digital divide between countries and contribute to global e-Government development.

MOU on e-Government Cooperation

Korea has signed MOUs with 45 countries and international organizations upon request for e-Government cooperation, and took part in various e-Government cooperation activities, such as high-level meetings with partner countries, cooperation projects, and dispatching delegations.

+82-2-2100-3957

e-Government Cooperation Center

Jointly operated for 3 years to find local e-Government needs and implement projects.

Operated in 7 countries so far, currently operating 1 in Indonesia and will open 2 new centers in Peru and Kenya.

+82-2-2100-3947

Korean e-Government Experience Program

Provide a customized training workshop on Korean e-Government to overseas public officials in charge of e-Government.

1,419 overseas public officials from 111 countries have participated in Korean e-Government experience program.

+82-2-2100-3955

e-Government Consulting

Provide consulting services on establishing informatization strategies and master plans based on Korea's knowledge and experience on e-Government.

+82-2-2100-3953

KOICA

Definitions: AI vs IA

(Source: The presentation by Dr. Jim Spohrer (IBM) © IBM UPWard 2016)

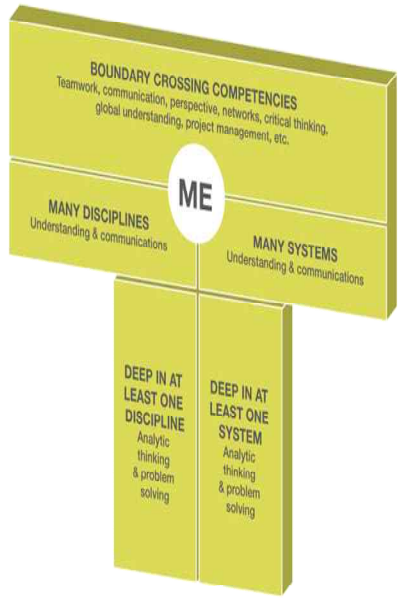
AI is Artificial Intelligence, or intelligence in machines (smart machines)

- **IA is Intelligence Augmentation, or people thinking and working together with smart machines.**
- **IA is what IBM calls "Cognitive Computing" and the smart machines are called "Watson Solutions" or more generally "Digital Cognitive Systems (Cogs)"**

A Fusion of Technologies with the Deep Learning Concept

Multi-disciplinary research: The Main Role

- Disciplines: Breadth of disciplines to tackle issues:
  - Psychology and cognitive science, philosophy, design and art, public policy and management, law and regulations
- Systems: Professional associations to tackle industry and system issues, including novice to expert progression on tasks
- Socio-technical system design loop and smart service systems



Types of AI

- **Tool**

- **Assistant**

- **Collaborator**

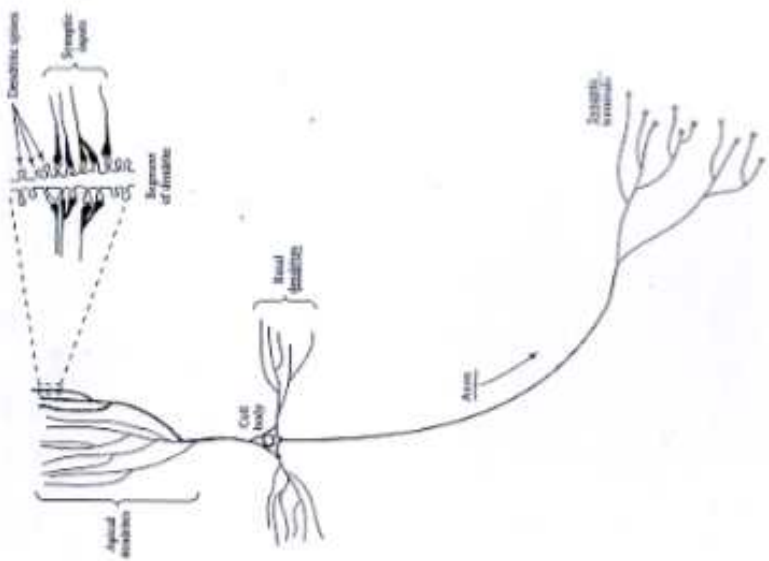
- **Coach**

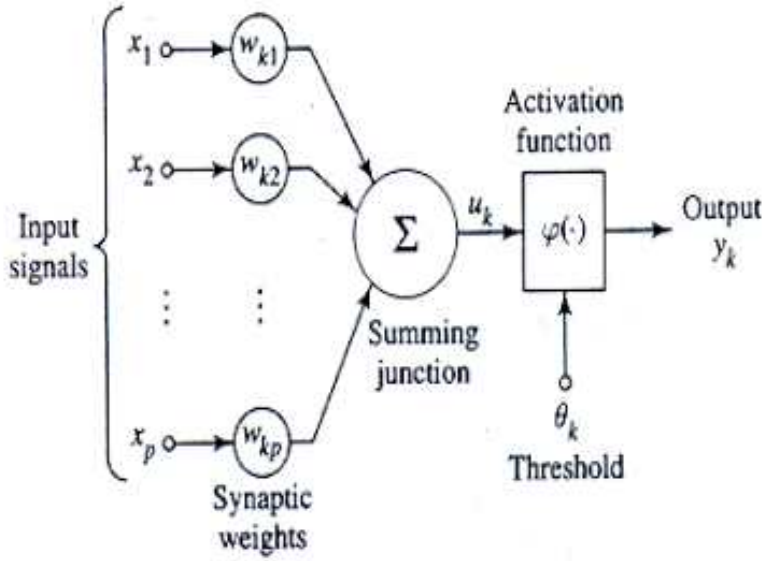
- **Mediator**

Types: Progression of models and capabilities

	Task & World Model/ Planning & Decisions	Self Model/ Capacity & Limits	User Model/ Episodic Memory	Institutions Model/ Trust & Social Acts
Tool	+	-	-	-
Assistant	++	+	-	-
Collaborator	+++	++	+	-
Coach	++++	+++	++	+
Mediator	+++++	++++	+++	++

Artificial Neural Networks

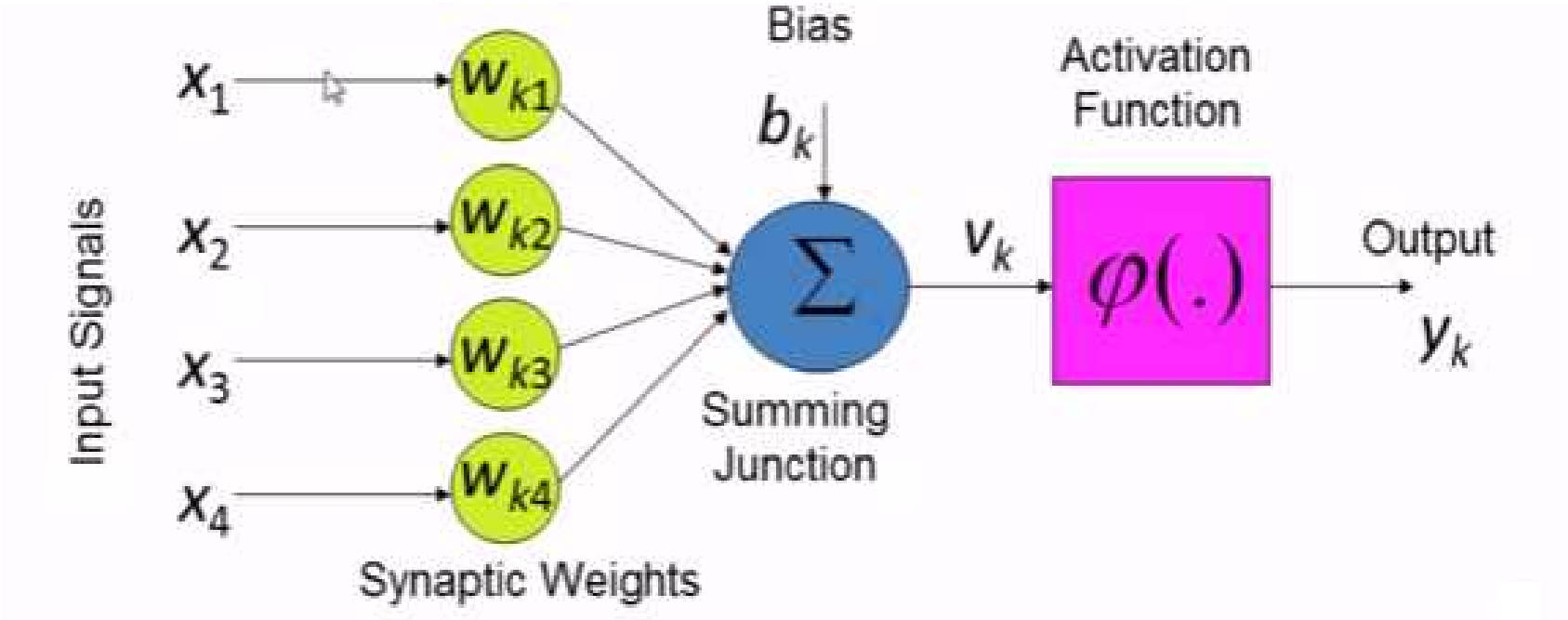




Dendrite **Soma** **Axon Spin**

- **Simple Perceptron**
- **The neuronal model we have just discussed is also known as a perceptron**
- **The perceptron is the simplest form of a neural network used for the classification of patterns said to be linearly separable**
- **Basically, it consists of a single neuron with adjustable synaptic weights and bias**

- **Simple Perceptron**



- In mathematical terms, a neuron k can be described by:

$$u_k = \sum_{j=1}^{m} w_{kj} x_j$$

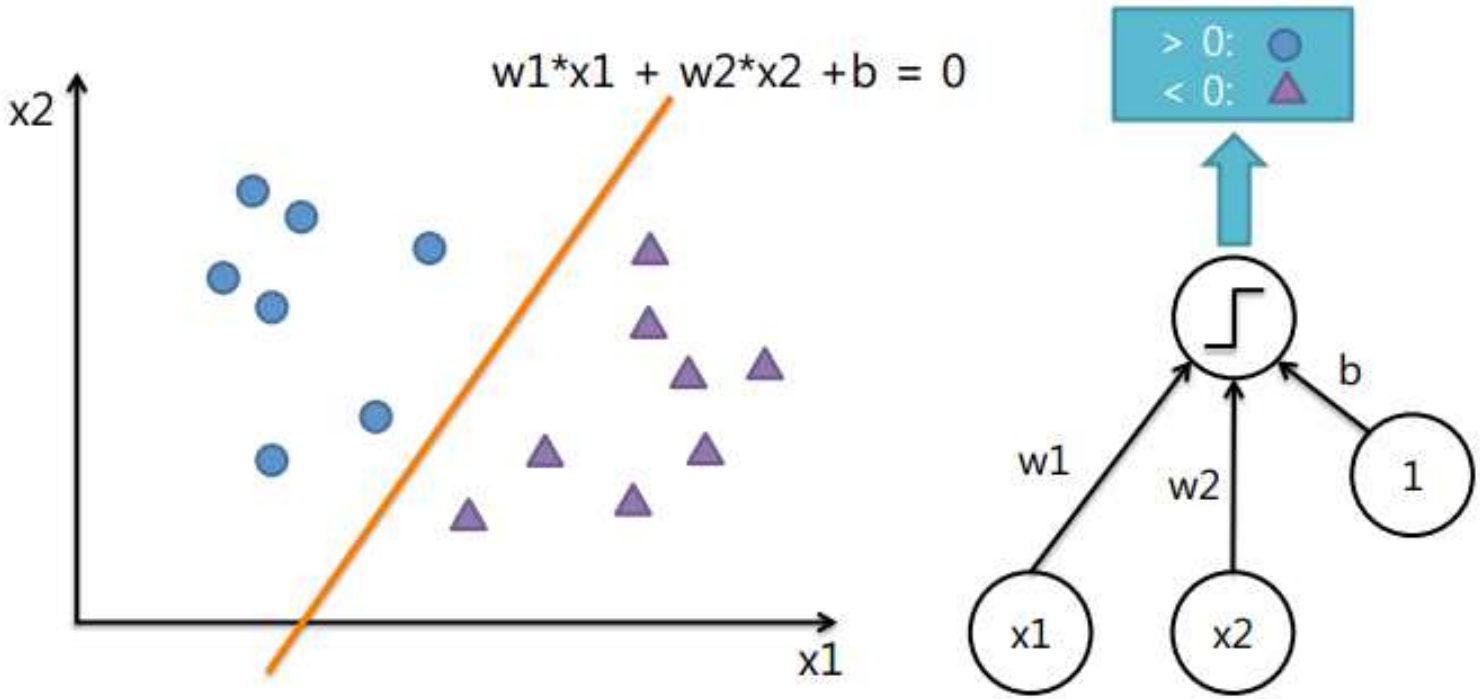
and

$$y_k = \varphi(u_k + b_k)$$

where u_k is the linear combiner output due to input signals.

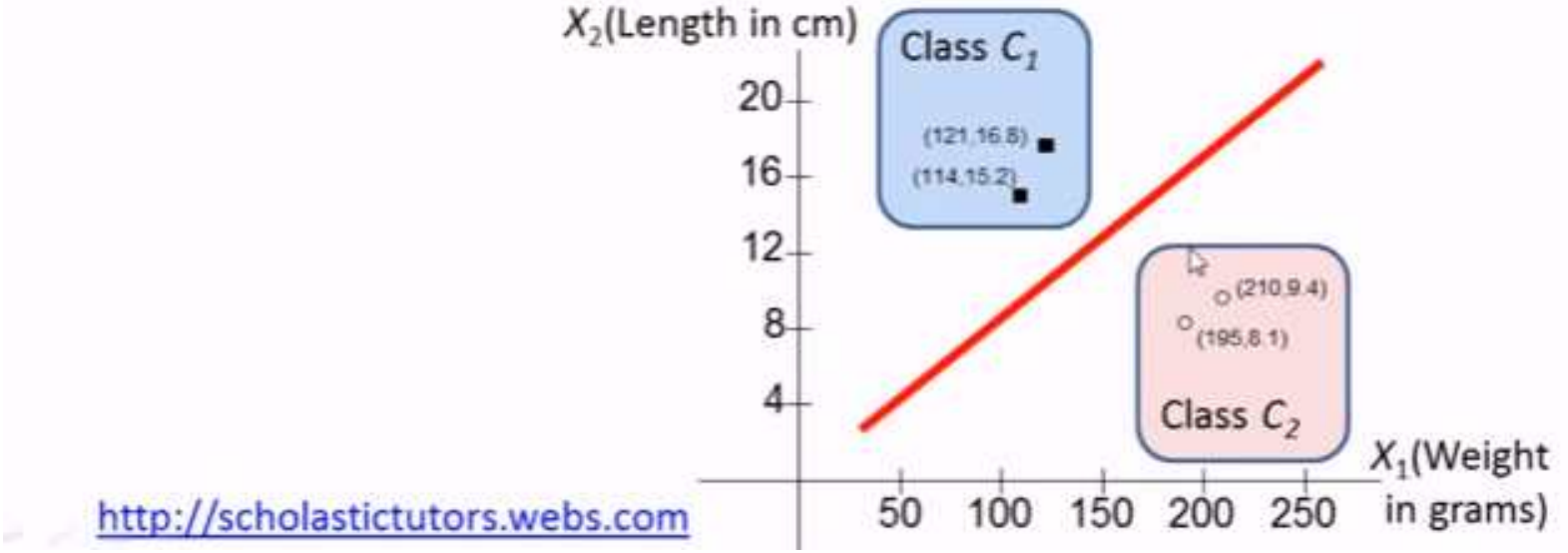
- Also

$$v_k = u_k + b_k$$


w1*x1 + w2*x2 +b = 0
x2
x1
> 0:
< 0:
b
w1
w2
1
x1
x2

Simple Perception – Example

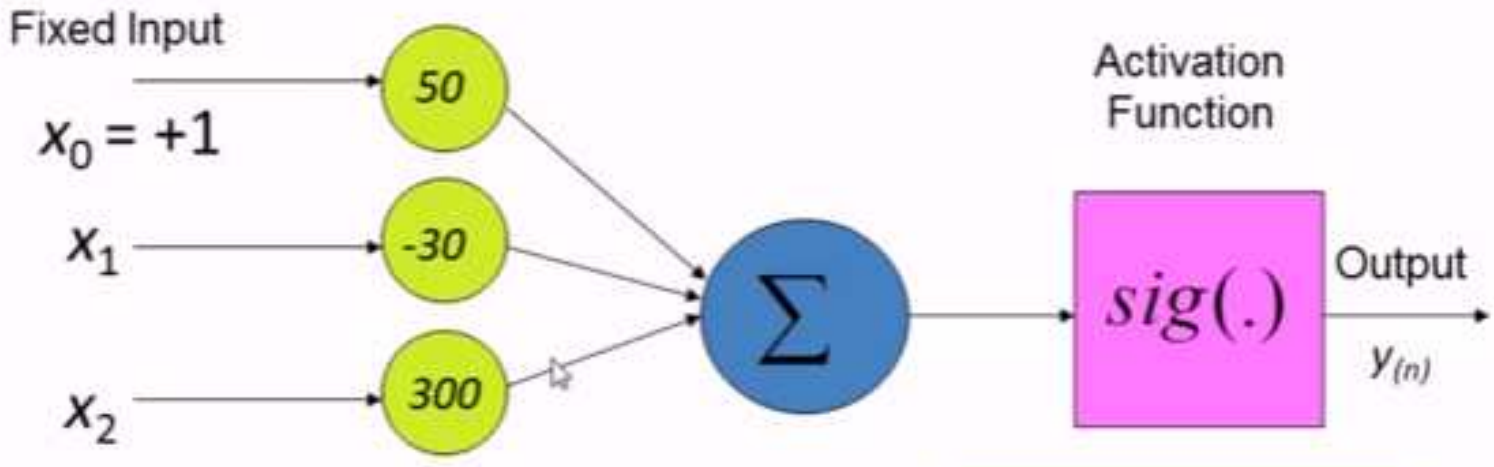
	Weight (grams)	Length (cm)
Fruit 1 (Class C1) **Banana**	121	16.8
	114	15.2
Fruit 2 (Class C2) **Apple**	210	9.4
	195	8.1



Simple Perception – Example

With correct initial weights and bias

$$\left.\begin{array}{l} w_1(0) = -30,\ w_2(0) = 300, \\ b(0) = 50,\ \eta = 0.01 \end{array}\right\} given$$



$$\text{sgn}(x) = \begin{cases} +1, & \text{if } x \geq 0 \\ -1, & \text{if } x < 0 \end{cases}$$

$$\left.\begin{array}{l} w_1(0) = -30,\ w_2(0) = 300, \\ b(0) = 50,\ \eta = 0.01 \end{array}\right\} \text{given}$$

Therefore the Initial Decision Boundary for this example is:

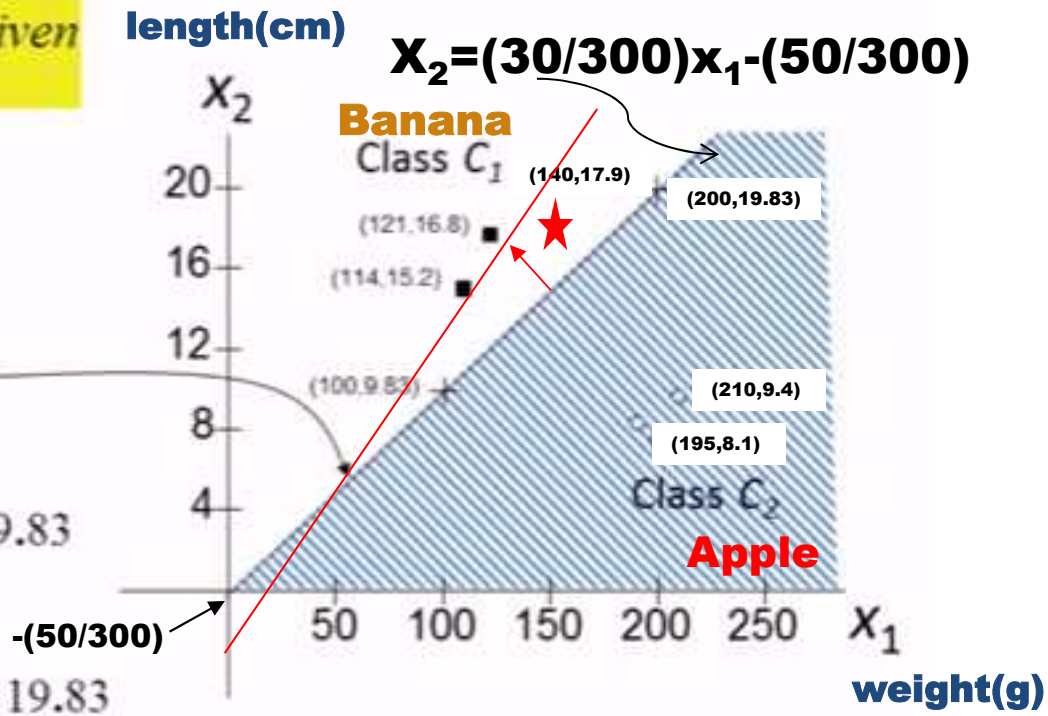
$$w_1 x_1 + w_2 x_2 + b = 0$$

$$-30x_1 + 300x_2 + 50 = 0$$

$$x_1 = 100,\ x_2 = \frac{30 \times 100 - 50}{300} = 9.83$$

$$x_1 = 200,\ x_2 = \frac{30 \times 200 - 50}{300} = 19.83$$

Initial hyper-plane does separate the two classes.



The computer adjusted the slope and reclassified ★ from Class C_1 to Class C_2 after learning based on Supervised Learning.

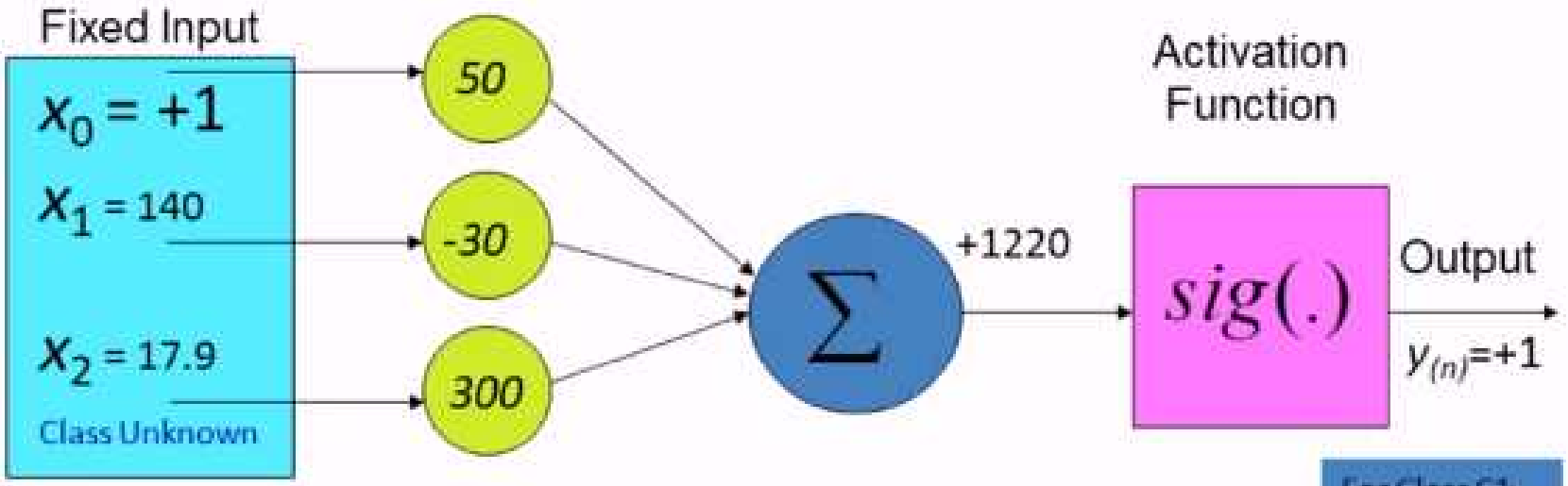
- **Simple Perception – Example**

-30*121 + 300*16.8 +50 = 1,460 > 0

→Class C_1

-30*210 + 300*9.4 +50 = -3,430 < 0

→Class C_2



Now use the above model to classify the unknown fruit.

$$\mathbf{x}(\text{unknown}) = [+1, 140, 17.9]^T$$

$$\mathbf{w}(3) = [50, -30, 300]^T$$

$$y(\text{unknown}) = \text{sgn}\left(\mathbf{w}^T(3)\mathbf{x}(\text{unknown})\right) = \text{sgn}(50 \times 1 - 30 \times 140 + 300 \times 17.9)$$

$$= \text{sgn}(1220) = +1$$

$\therefore$ this unknown fruit belongs to the class C_1.

Internet of Things (IoT)

- **Extending the current Internet and providing connection, communication, and inter-networking between devices and physical objects, or "Things," is a growing trend that is often referred to as the *Internet of Things*.**

- **“The technologies and solutions that enable integration of real world data and services into the current information networking technologies are often described under the umbrella term of the Internet of Things (IoT)”**

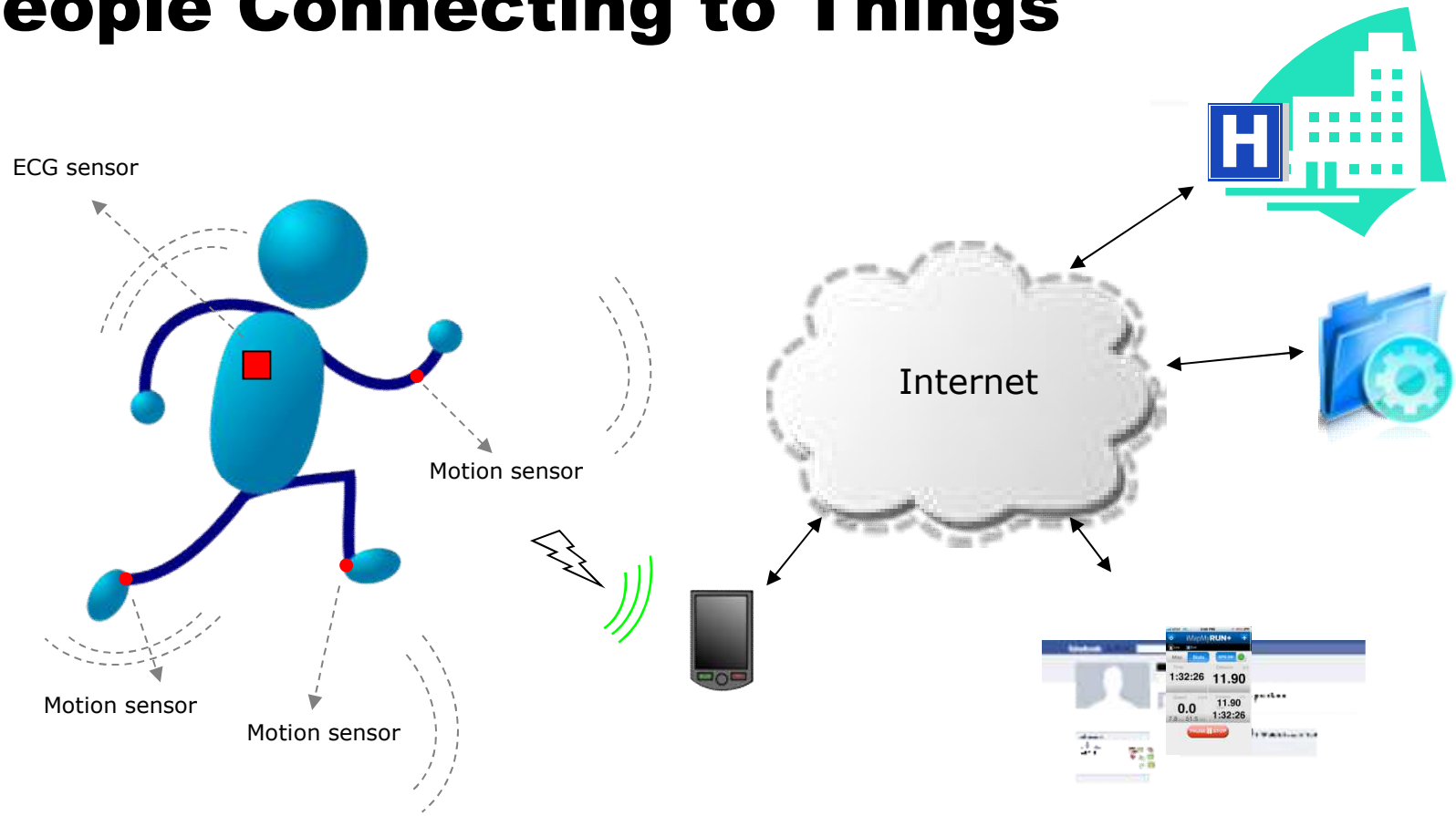
Sensor devices are becoming widely available

- Programmable devices
- Off-the-shelf gadgets/tools

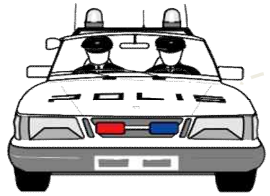
More “Things” are being connected

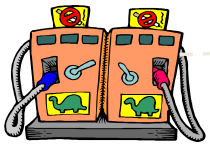
Home/daily-life devices
Business and
Public infrastructure
Health-care
…

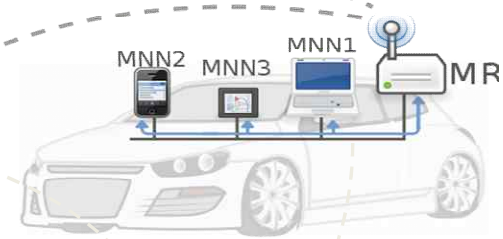
People Connecting to Things



Things Connecting to Things







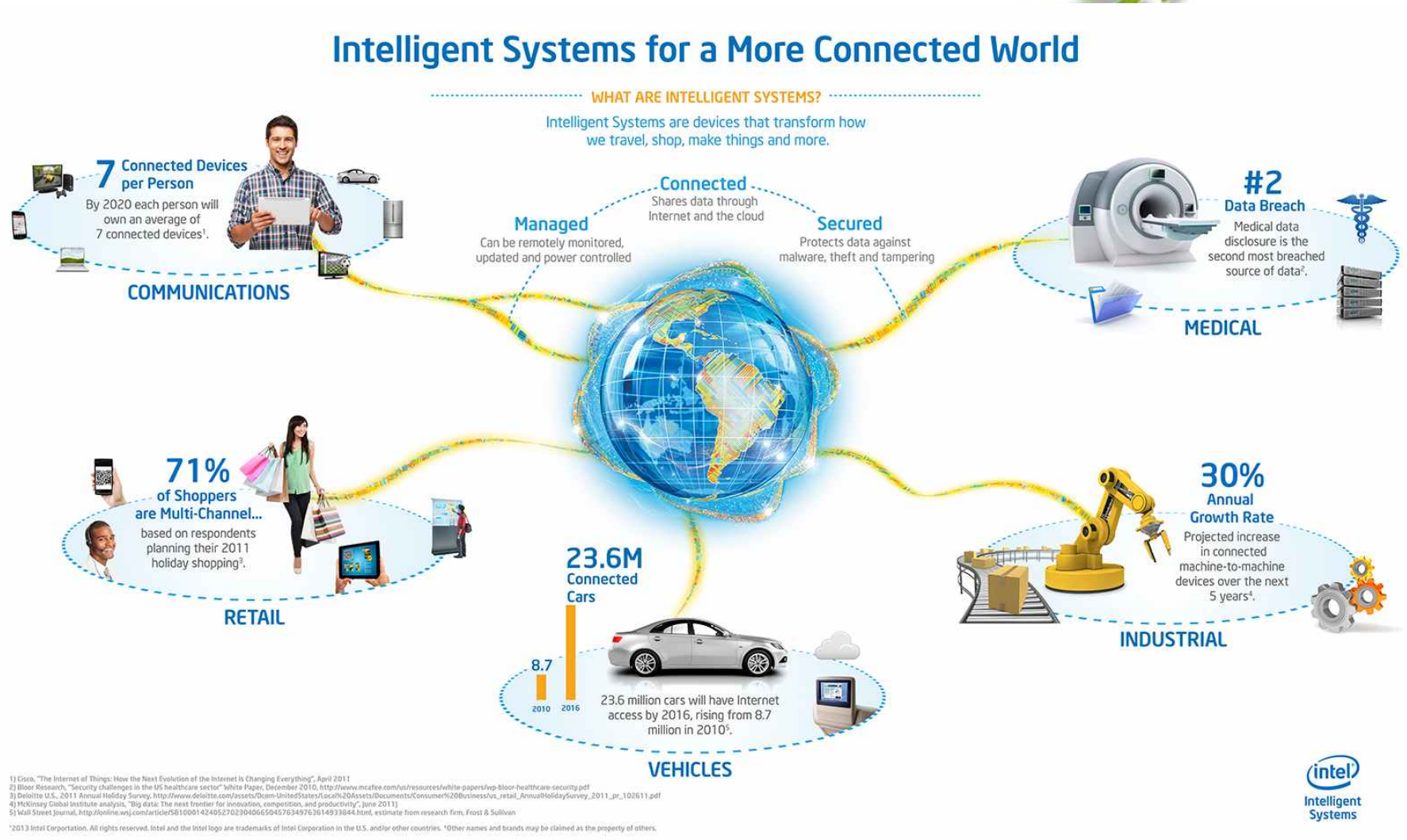
- Complex and heterogeneous resources and networks

How are the networks changing?

- **Extensions**
  - **More nodes, more connections, IPv6, 6LowPan,…**
  - **Any TIME, Any PLACE, Any THING**
  - **M2M, IoT**
    - **Billions of interconnected devices,**
    - **Everybody connected.**
- **Expansions**
  - **Broadband**
- **Enhancements**
  - **Smart networks**
  - **Data-centric and content-oriented networking**
  - **Context-aware (autonomous) systems**

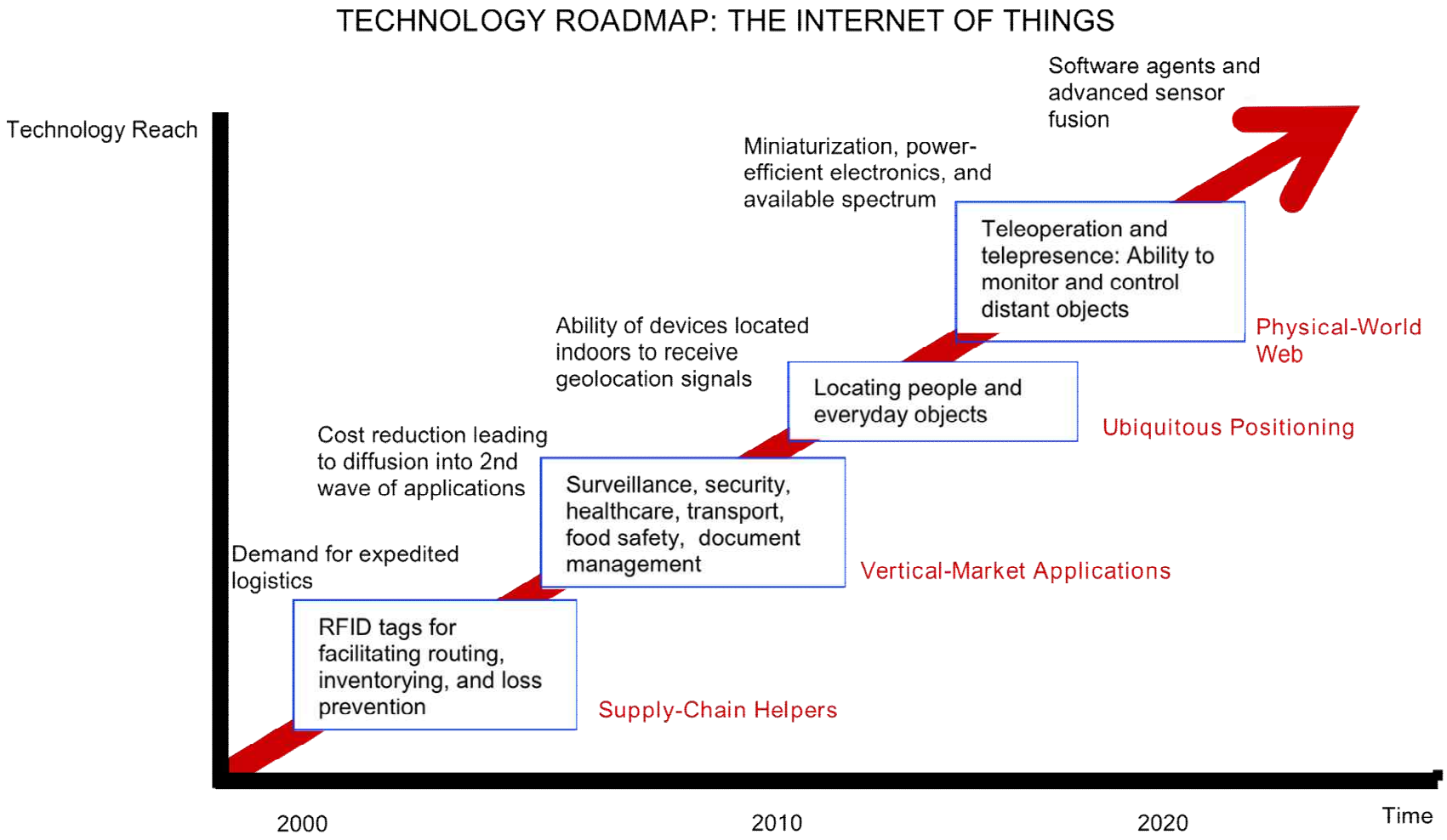
Why should new professionals learn about IoT?

- **Business Trends**
- **Emerging Technologies**
- **Growing IoT Services and Application**
- **Innovating Clients' Companies through New Technologies**



Source: http://blog.trentonsystems.com/internet-of-things-crosses-business-personal-boundaries/

Technology trend



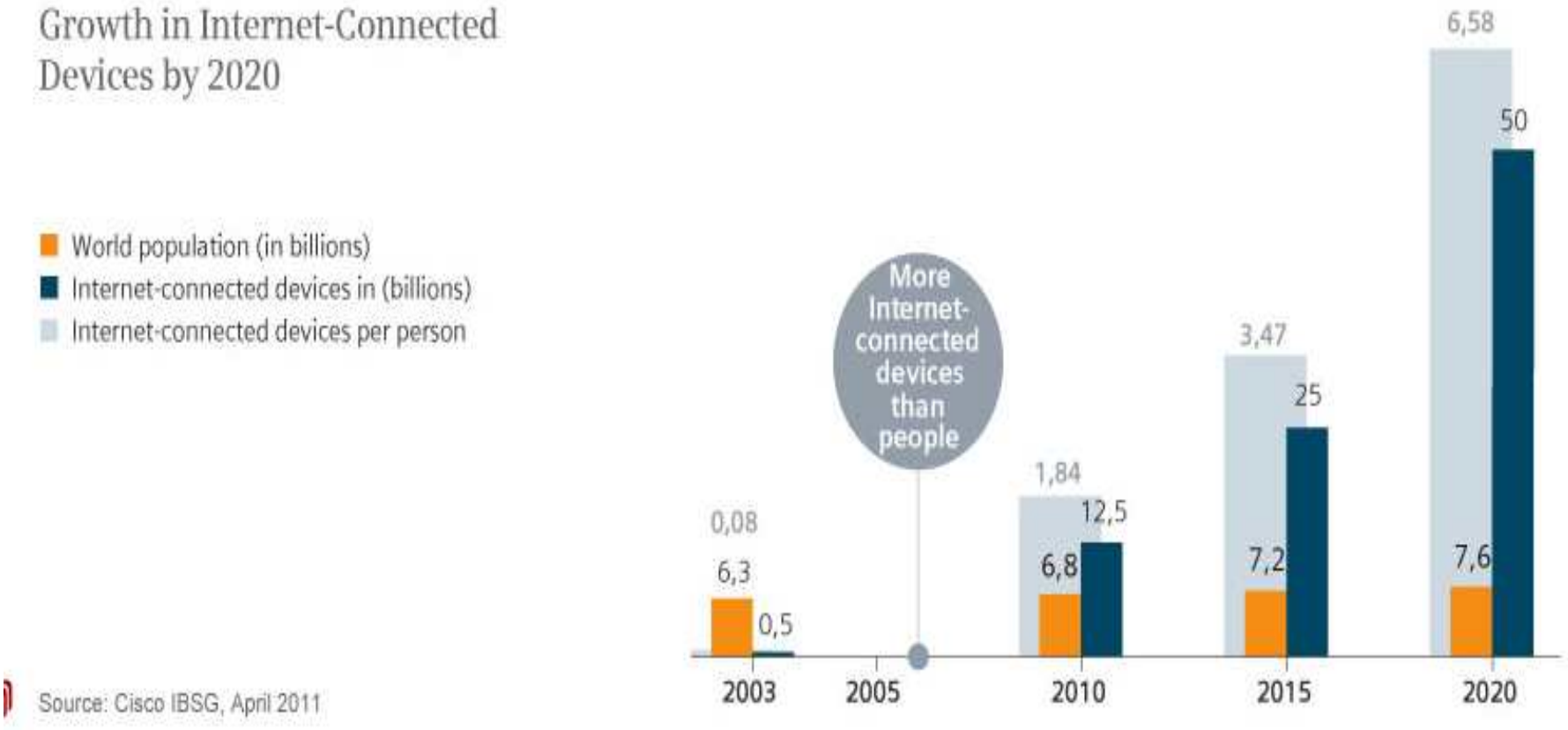
Source: SRI Consulting Business Intelligence

Market growth

- "According to a study conducted by Frost & Sullivan in 2011, the global RFID market of $3 billion to $4 billion (in 2009) will grow by twelve percent per year through 2016 and reach a volume of approximately $6.5 billion to almost $9 billion."
- 80 percent of all households in the European Union are expected to have intelligent power meters by 2020.
- A building's energy management can then be monitored and administered remotely via a smartphone or a PC. Market experts predict that this global market, which represented $5.3 billion in 2010.
- In February 2012 the Chinese government therefore decided to set up a fund of approximately $775 million to support this field in the next five years. It will grow to $11 billion by 2015.
  - This sector is expected to grow to $116 billion by 2015, according to a report published by the Xinhua News Agency in late 2010.

Source: Siemens, http://www.siemens.com/innovation/apps/pof_microsite/_pof-fall-2012/_html_en/facts-and-forecasts-growth-market-of-the-future.html

Internet Connected devices



Source: Siemens, http://www.siemens.com/innovation/apps/pof_microsite/_pof-fall-2012/_html_en/facts-and-forecasts-growth-market-of-the-future.html

Internet of Things Module

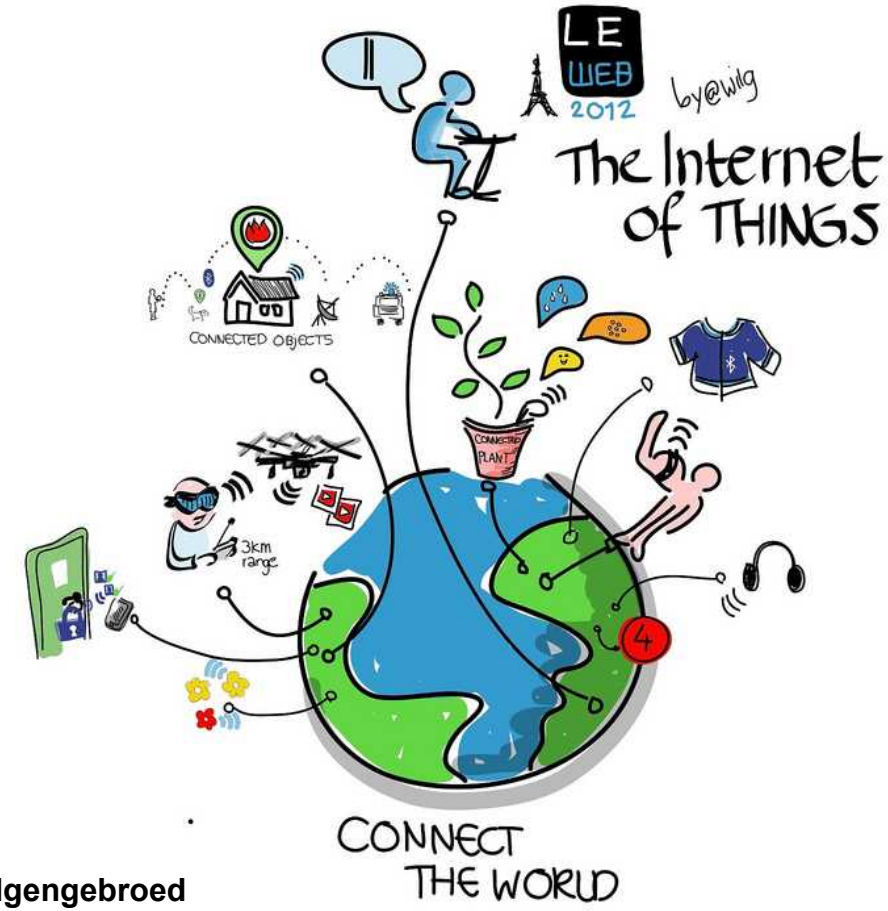


Image courtesy: Wilgengebroed

- Today, we are at the beginning of a Fourth Industrial Revolution. Developments in genetics, artificial intelligence, robotics, nanotechnology, 3D printing and biotechnology, to name just a few, are all building on and amplifying one another.
- Smart systems—homes, factories, farms, grids or cities—will help tackle problems ranging from supply chain management to climate change.
- The rise of the sharing economy will allow people to monetize everything from their empty house to their car.

- The Future of Jobs Report is a first step in becoming specific about the changes at hand. It taps into the knowledge of those who are best placed to observe the dynamics of workforces—Chief Human Resources and Strategy Officers—by asking them what the current shifts mean, specifically for employment, skills and recruitment across industries and geographies.

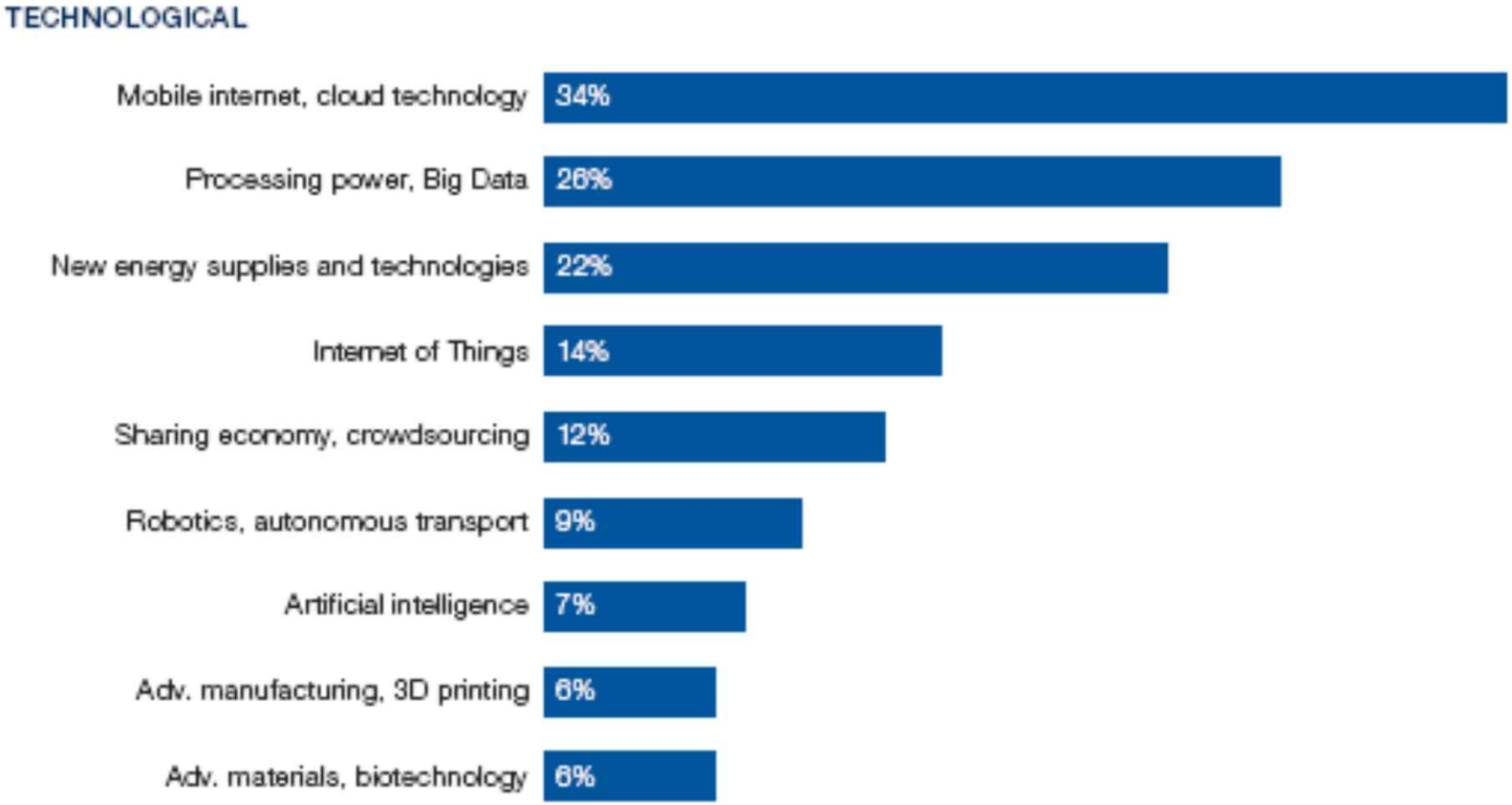
- The question, then, is how professionals will react to these developments.
- To prevent a worst-case scenario—technological change accompanied by talent shortages, mass unemployment and growing inequality—reskilling and upskilling of today's workers will be critical.

- In particular, business collaboration within industries to create larger pools of skilled talent will become indispensable, as will multi-sector skilling partnerships that leverage the very same collaborative models that underpin many of the technology-driven business changes underway today.

- To ensure that we achieve this vision, we must become more specific and much faster in understanding the changes underway and cognizant of our collective responsibility to lead our businesses and communities through this transformative moment.

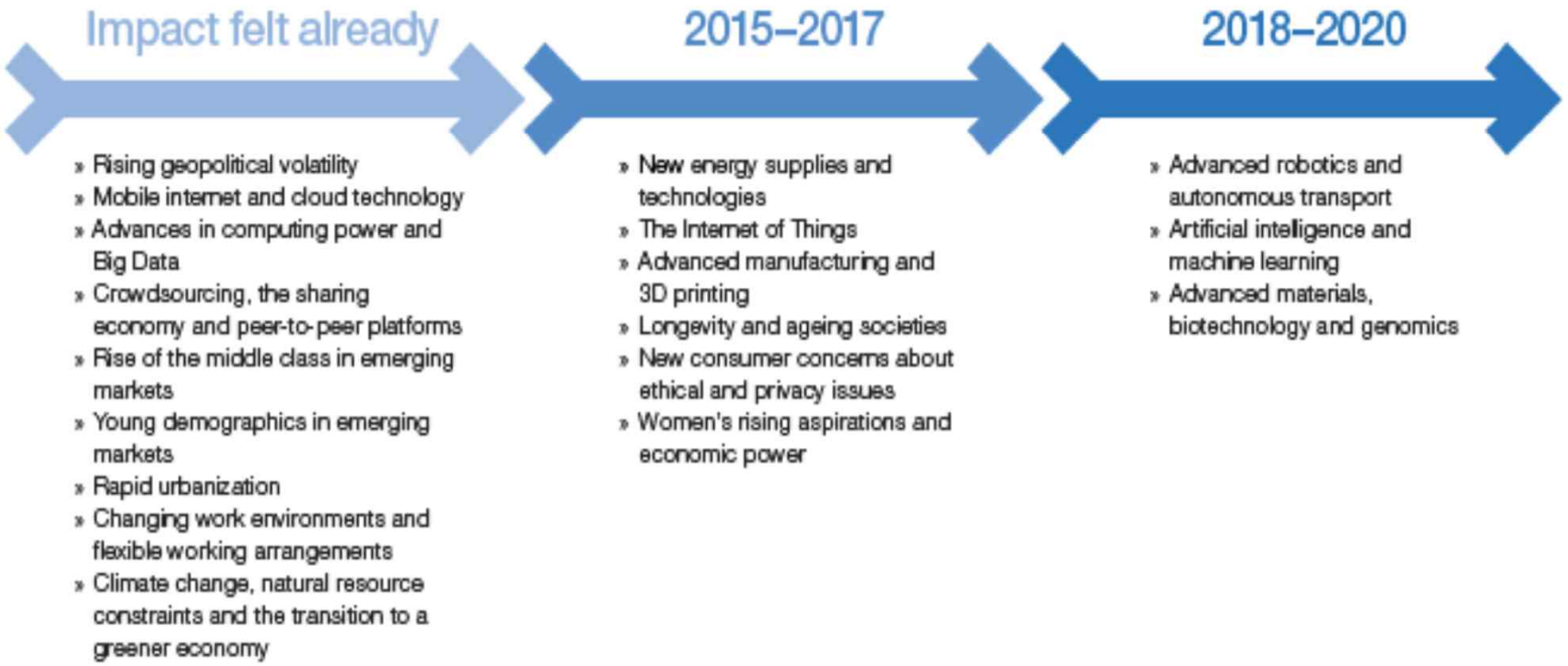
- Disruptive changes to business models will have a profound impact on the employment landscape over the coming years.
- Many of the major drivers of transformation currently affecting global industries are expected to have a significant impact on jobs, ranging from significant job creation to job displacement, and from heightened labor productivity to widening skills gaps.

- By one popular estimate, 65% of children entering primary school today will ultimately end up working in completely new job types that don't yet exist.
- In such a rapidly evolving employment landscape, the ability to anticipate and prepare for future skills requirements, job content and the aggregate effect on employment is increasingly critical for businesses, governments and individuals in order to fully seize the opportunities presented by these trends—and to mitigate undesirable outcomes.



Source: Future of Jobs Survey, World Economic Forum.
Note: Names of drivers have been abbreviated to ensure legibility.

Timeframe to impact industries, business models

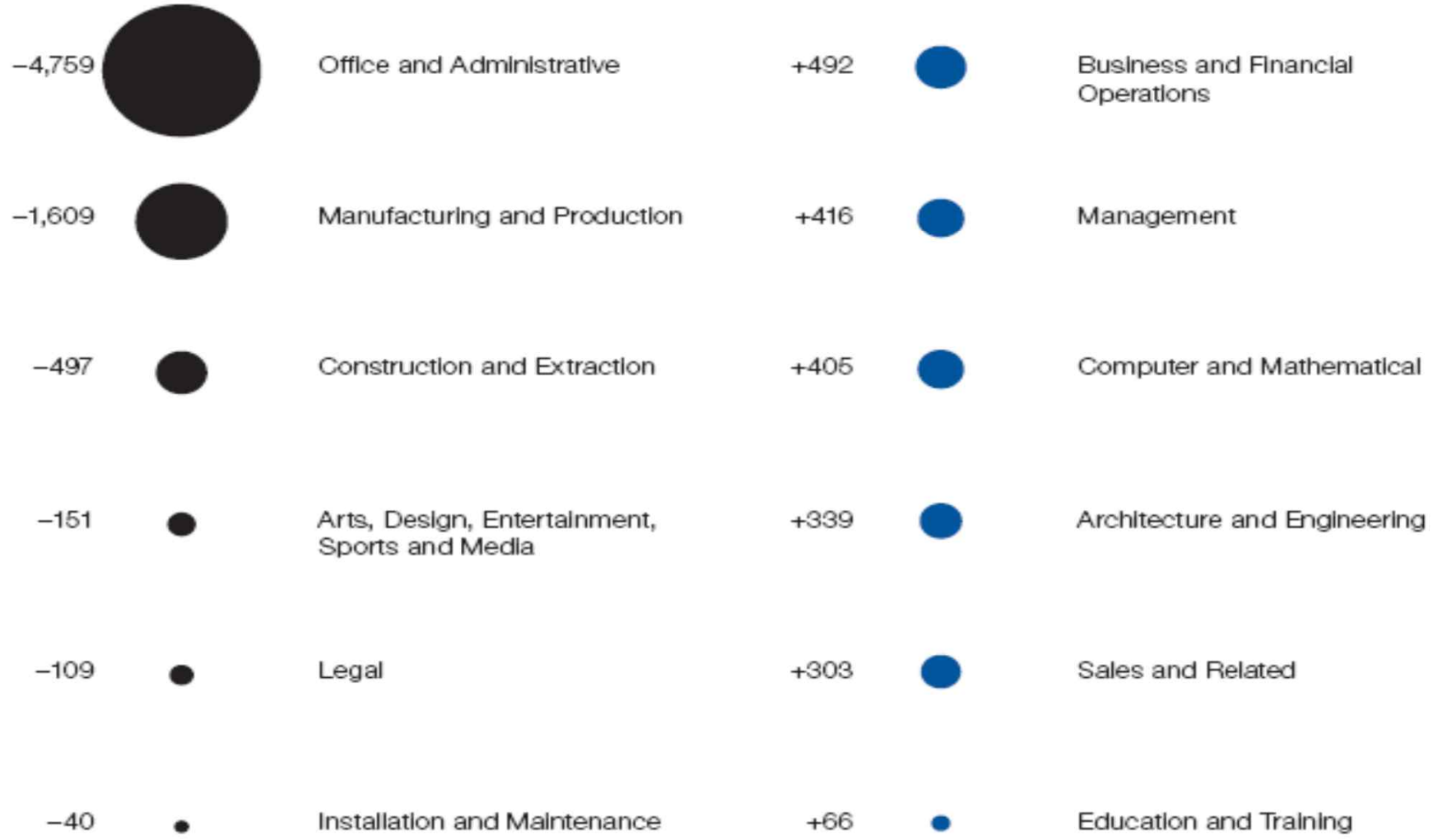


Impact of Disruptive Change on Employment

- By contrast, further unpacking the bundle of technological drivers of change in the mold of the Fourth Industrial Revolution yields a rather more optimistic picture regarding the job creation potential of technologies such as Big Data analytics, mobile internet, the Internet of Things and robotics.

Net employment outlook by job family, 2015–2020

Employees (thousands, all focus countries)



Cross-functional Skills

Social Skills
- Coordinating with Others
- Emotional Intelligence
- Negotiation
- Persuasion
- Service Orientation
- Training and Teaching Others

Systems Skills
- Judgement and Decision-making
- Systems Analysis

Complex Problem Solving Skills
- Complex Problem Solving

Resource Management Skills
- Management of Financial Resources
- Management of Material Resources
- People Management
- Time Management

Technical Skills
- Equipment Maintenance and Repair
- Equipment Operation and Control
- Programming
- Quality Control
- Technology and User Experience Design
- Troubleshooting

V. A Case of Applying Big Data Tools to Strategic Decision Making in a Private Company

(Source: Kyeong Seok HAN and Seok Yong Yun, A Study on Establishing Competitive Advantage Strategies based on Patent Data Investigation using TF-IDF(Term Frequency – Inverse Document Frequency) and Network Analysis, GLOBAL CONFERENCE ON BUSINESS MANAGEMENT AND ECONOMICS 2018)

V-I. Introduction

- This research paper explores the value of data that were considered only raw data for information and knowledge and were not fully recognized 10 years age.
- However, many companies are still using the data analysis as “Descriptive and Diagnostic Analysis” rather than “Predictive and Strategic Analysis” in these days.
- This research suggests how to establish strategies practically to achieve the competitive advantage based on the big data analysis and machine learning tools (Jun Sunghae(2013)).

- We need to analyze the data about the competitors to establish strategies to achieve the competitive advantage, but it is very difficult to get the current data about them. However, the patent registration data about the competitors are open to the public while it is protected legally for 20 years.
- This research will show how to establish strategies against the competitive companies based on the analysis of competitors' technological strategies using the quantitative and qualitative patent data that are open to the public (Max H. Boisot, Ian C. MacMillan, Kyeong Seok Han (2008).

V-II. Theoretical Background

Patent Data

- Patent systems have the regulations to protect the right of patent holders to promote the invention for the national development.
- Once the patent passed the evaluation processes, it will be open to the public after one and half years and its right will be exclusively protected for 20 years legally (KIPO(Korean Intellectual Property Office), 2018).

Structure of International Patent Classification.

- International Patent Classification Patent codes were implemented in 1967 with about 70,000 codes and shared by about 180 member countries. As shown in [Table 1] IPC codes consist of Section, Class, Subclass, Main Group and Subgroup. [Table 2] shows that each section includes classified technologies. This scheme is an international standard, even though the system is not perfect.

Network Analysis

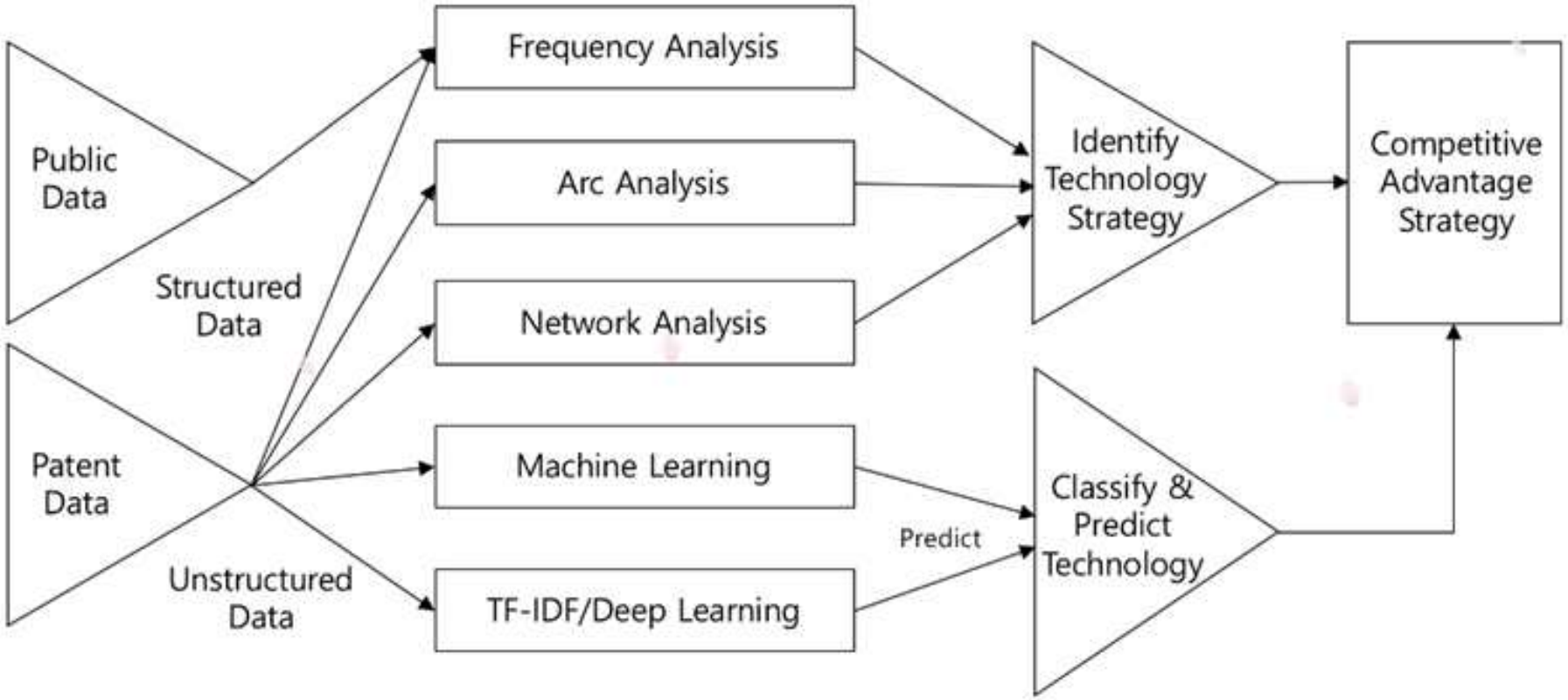
- Network Centrality Theory. The theory is used to identify the importance of a node that is influential in the network.
- The theory includes three indices i.e. Degree Centrality, Betweenness Centrality, Closeness Centrality (Japanese Ministry of General Affairs(2012). We can identify the strategies and core technologies of the competitors using the network analysis.

Establishing Strategies for Competitive Advantage Based on Competitor's Patent Analysis

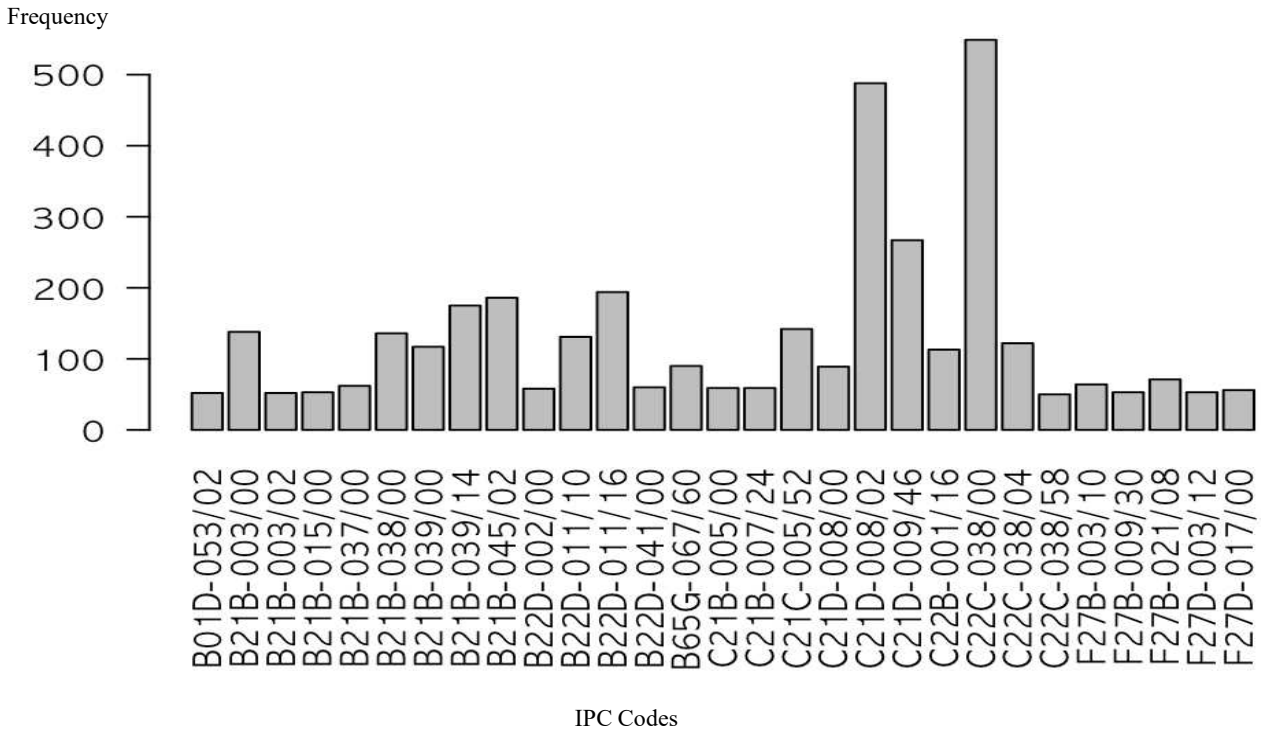
- This research uses 5311 patent data of Company B collected from 2000 to 2014 using KIPRIS(Korean Intellectual Property Rights Information Service).
- Company B is a competitor of Company A which wants to establish strategies for competitive advantage based on competitor's patent analysis.

Establishing Strategies for Competitive Advantage Based on Competitor's Patent Analysis

- The data consist of 55 fields such as patent number, name, IPC(International Patent Classification) code, patent applicant, inventor, application date, registration date, open date, international patent application number, international patent application date, abstract, application items, family data, and so on.
- We collected more data from KISTI(Korea Institute of Science and Technology Information), KIPO(Korean Intellectual Property Office), Google, etc.



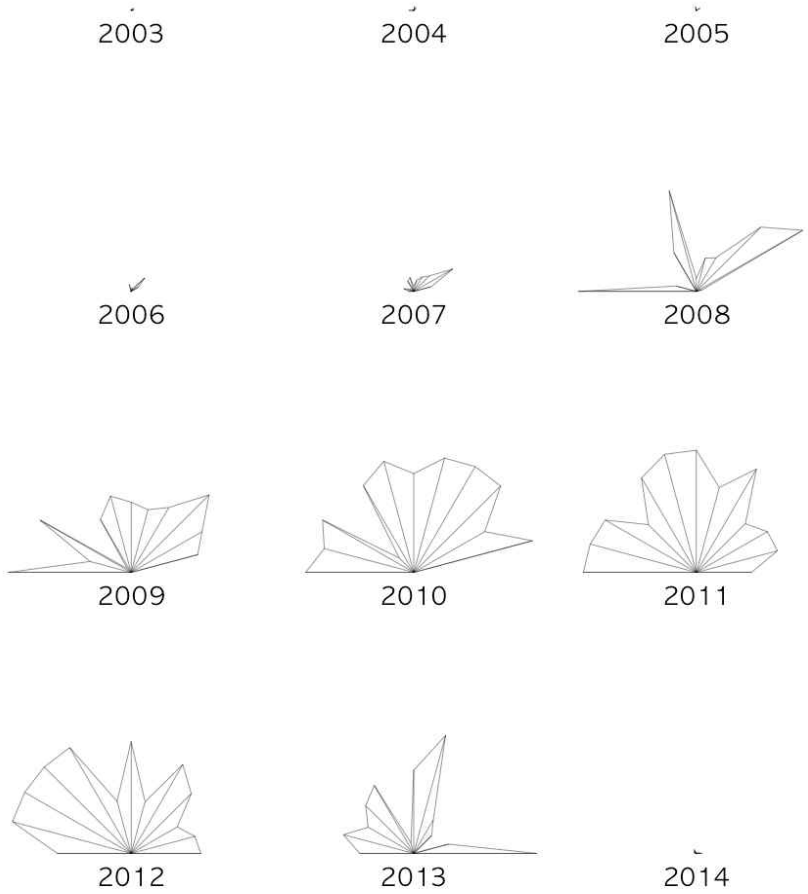
[Figure 1] A Research Model to Establish Strategies for Competitive Advantage Based on Competitor's Patent Analysis-PATS(Patent Analytics for Technology Strategy) Model



[Figure 2] A frequency analysis of IPC codes with more than 50 patent registrations

Plot Analysis

We performed the mosaic plot analysis to find the diversity of patents using yearly-based data. [Figure 3] shows that the numbers and types of patents has increased while the growth of company was realized during the periods.

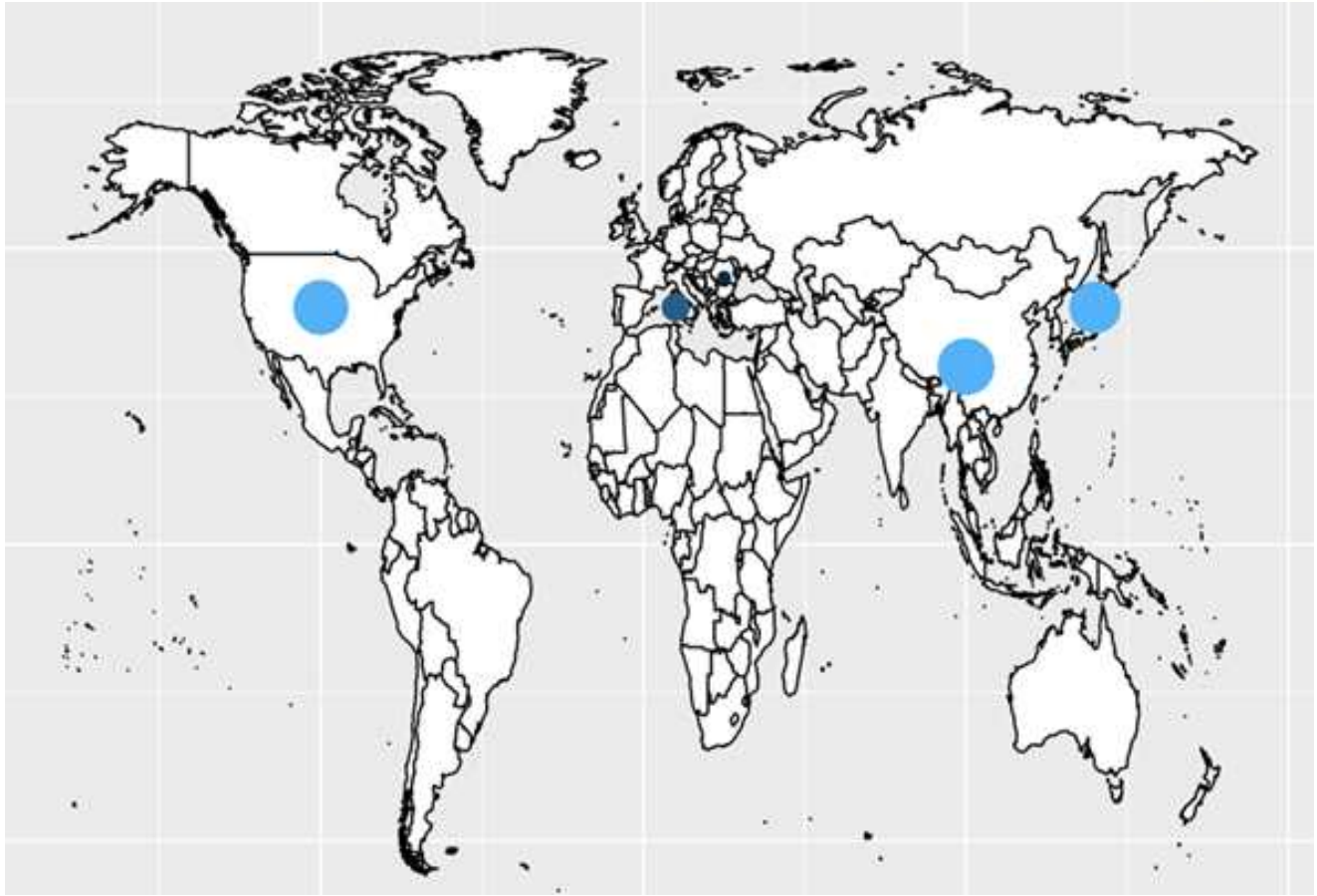


[Figure 3] Mosaic plots of IPC codes with yearly-based patent registration numbers

Word Cloud Analysis

- Word cloud analysis based on the patent names allows us to understand the competitor's technology trends.
- In the earlier stage the competitor looks interested in 'Rolled Steel' like technologies, but in the later stage it is interested in 'Apparatus' like technologies as shown in [Figure 4].

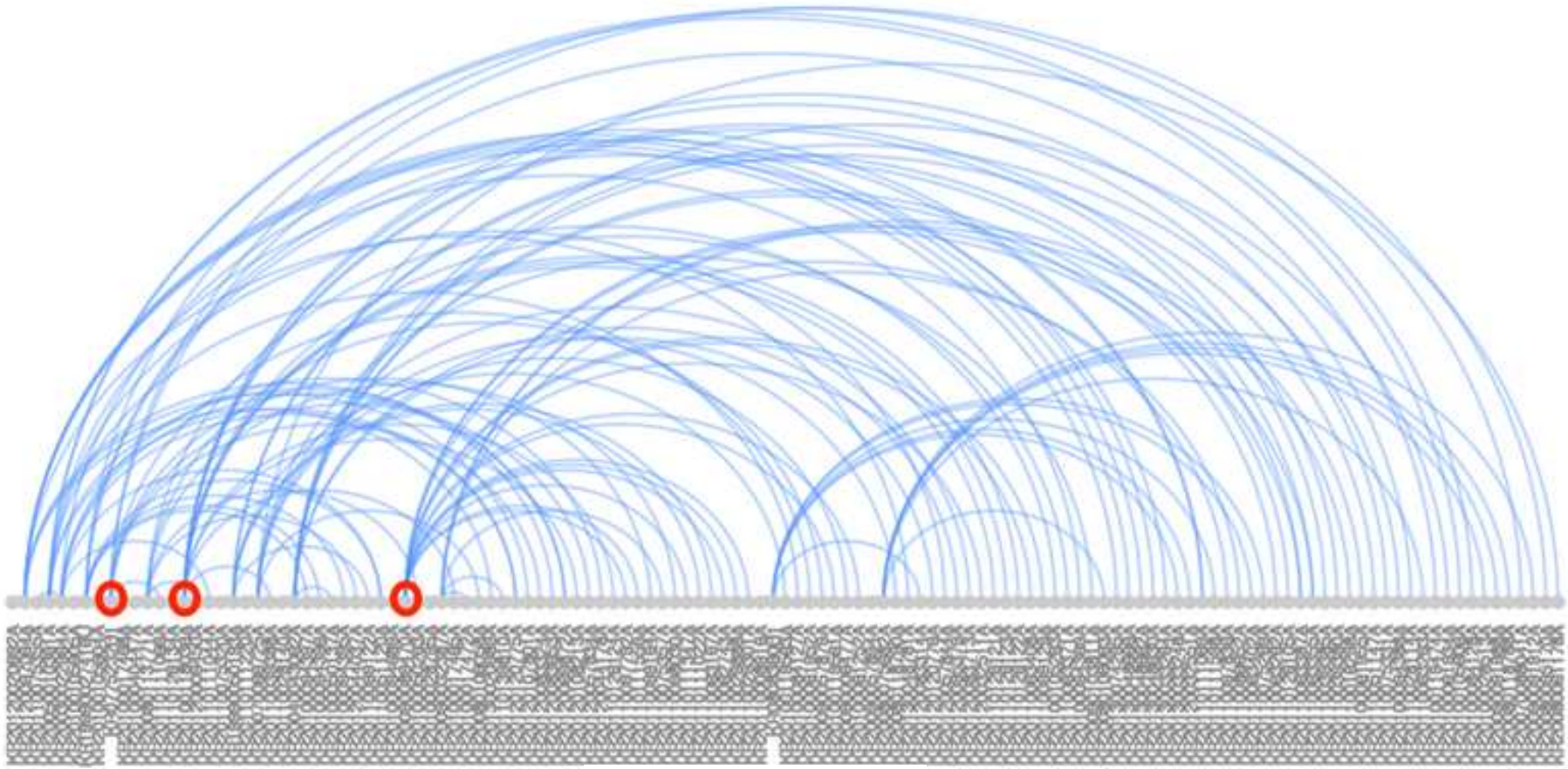
[Figure 4] The "Word Cloud" based on yearly-based patent registration names



[Figure 5] The geographical distribution of patent registration of the competitor

Arc Analysis

- The key patent can be identified through the IPC reference analysis. The competitor, Company B focused on sintering processes and the fuel raw material equipment as shown in [Figure 6].
- [Figure 6] also shows that most referred patents are KR2009008***** with 15 times reference and KR2010011***** with 13 times reference.

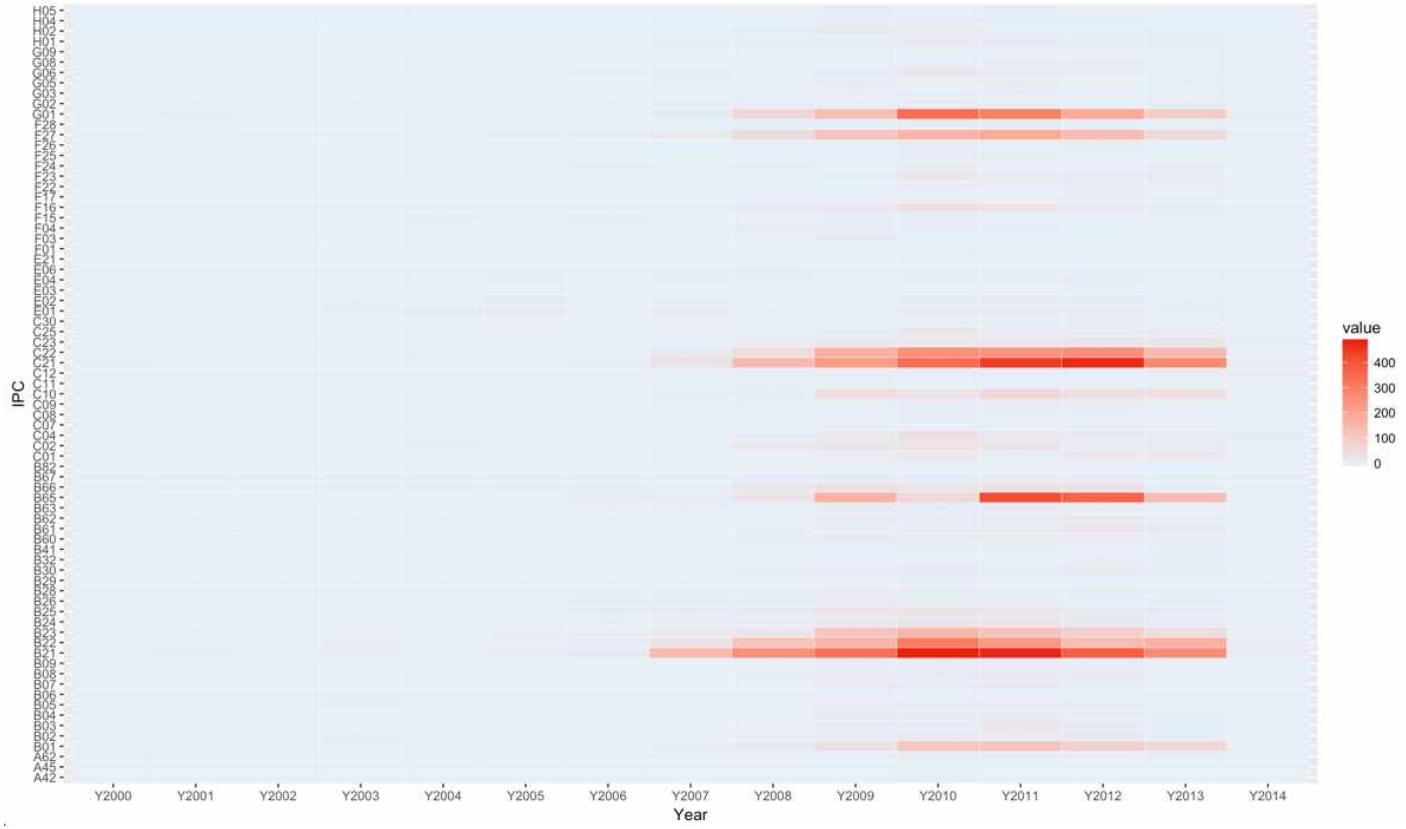


Registered IPCs

[Figure 6] The arc analysis based on registered IPC

Heatmap Analysis

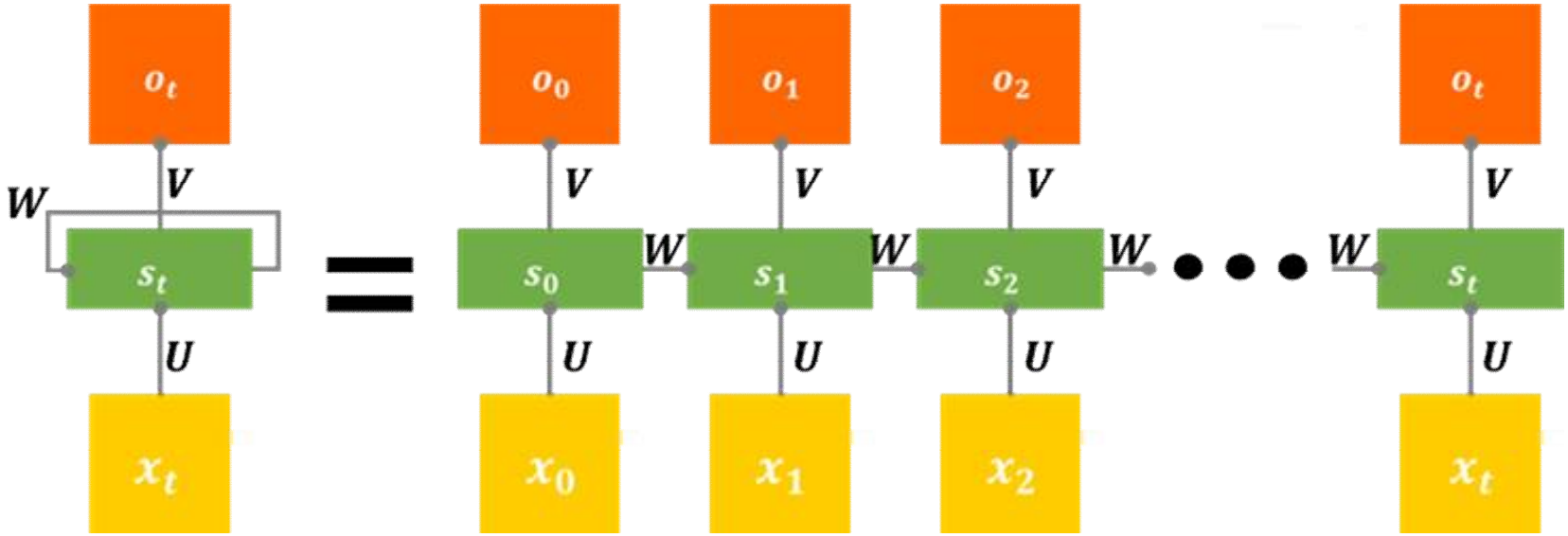
If we draw the frequencies of competitor's patents on each year,
we can find that the competitor's patents are developed during certain periods
as shown in [Figure 7].



[Figure 7] The yearly based frequency(Heatmap) of competitor's IPC codes

- RNN(Recurrent Neural Network) is a method of the deep learning based on neural network. RNN is an applied model evolved from DNN(Deep Neural Network).
- We want to forecast the competitor's technology strategies based on competitor's patent data and RNN method. [Figure 8] is an example of RNN model.

V. Technology Prediction Based on LSTM (Long Short-Term Memory)



where, x_t are input values at time t

s_t are hidden layer statuses at time t

o_t are output values at time t

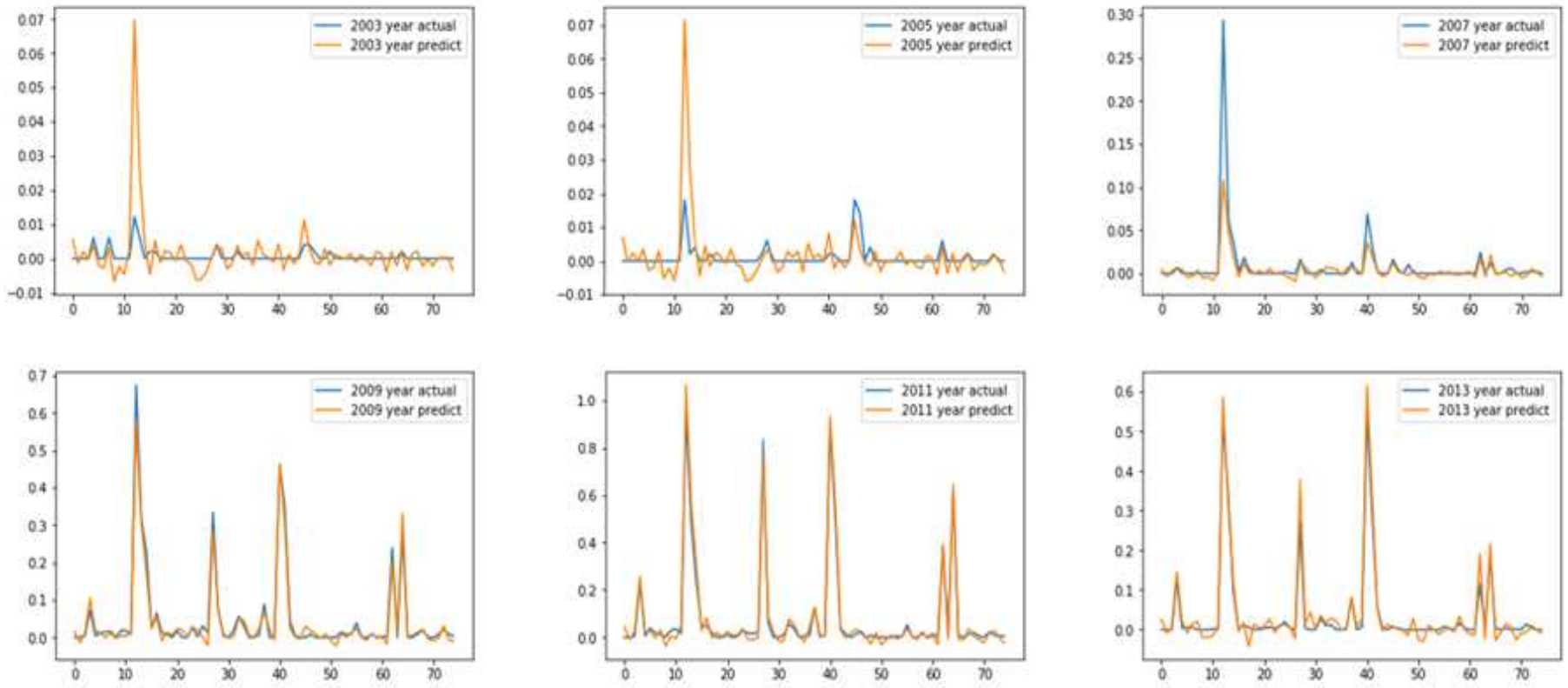
U, V, W are weight values attained after machine learning

[Figure 8] An example of RNN model

V. Technology Prediction Based on LSTM (Long Short-Term Memory)

- However, RNN has a vanishing gradient problem and we used LSTM(Long Short-Term Memory) model that is an improved RNN model. The forecasting results per year are shown in [Figure 9].
- In the graph x-axis represents IPC codes and y-axis represents IPC registration frequencies. The accuracy is 81.82%, which is very accurate.

V. Technology Prediction Based on LSTM (Long Short-Term Memory)



[Figure 9] The forecasting accuracy of LSTM model

Main Process	Sub Process	Code	Keyword
Iron Making	Sintering	0101	Iron Ore, Subsidiary Raw Material, Sized Ore, Sinter, Limestone
	Coke	0102	Coke, Bituminous Coal, Coal, Powdered Coal, Pulverized Coal
	Subsidiary Products	0103	COG, BOG, Gas, Crude Light Oil, Cement
	Blast Furnace	0104	Pig Iron, Torpedo Car
Steel Making	Refining	0201	Iron, Scrap, Subsidiary Raw Material, Oxygen, Conventional Blowing, Argon, Nitrogen
	Refinement	0202	RH, Degassing, Inclusion, LF, Desulfur, Powder
	Continuous Casting	0203	Turndish, Mold, Cooling, Slab, Bloom, Billet
Rolling	Hot Rolling	0301	Reheating Furnace, Roughing Mill, Roll, Cooler, Winding, Hot Coil
	Cold Rolling	0302	Automobile, Annealing, Galvanize, Plating, Steel Sheet, Directional Properties

[Table 4] Classified processes and the keywords

VI. Automatic Classification of Competitor's Patent Data Based on TF-IDF

Patent Number	Iron Making				Steel Making			Rolling		Domain Expert Judgment
	Sintering	Coke	Subsidiary Products	Blast Furnace	Refining	Refinement	Continuous Casting	Hot Rolling	Cold Rolling	
KR2012006*****	100	200	0	200	300	1365	100	200	0	Refinement
KR2011006*****	36	26	203	28	100	110	200	300	100	Subsidiary Products
KR2011001*****	0	0	0	0	0	10	987	0	0	Continuous Casting
KR2010010*****	0	0	0	74	0	0	1580	100	0	Continuous Casting
KR2009007*****	0	1007	100	100	0	0	0	0	0	Coke
KR2009010*****	0	106	106	106	10	110	300	728	100	Hot Rolling

[Table 5] The part of patent classification of the competitor calculated by TF-IDF (Term Frequency – Inverse Document Frequency) statistics

Future Research

- This research paper shows that we can establish strategies practically to achieve the competitive advantage against our competitor based on the big data analysis and machine learning tools.

Future Research

- In other words, we showed how to establish strategies against the competitive companies based on the analysis of competitors' technological strategies using the quantitative and qualitative patent data that are open to the public using Frequency Analysis, Arc Analysis, Network Analysis, Heatmap Analysis, TF-IDF, LSTM and so on.
- In the future this research model will be able to be applied for other industries.

VI. Conclusion

- **e-Government professionals need to prepare the future public administration and business processes based on the Big Data-based Innovation through the New Body of Knowledge and Strategies to Adjust Ourselves to Disruptively Changing Environment in the Fourth Industrial Revolution Era.**

THANK YOU!

Q&A

Reference

- https://www.ft.com/content/9f0a8838-fa25-11e7-9b32-d7d59aace167, “Can big data revolutionize policymaking by governments?” in Make Big Data open, accessible and available in real time / From Gavin Hayman, Twickenham, Middx, UK Copyright The Financial Times Limited 2018.
- Zaher Ali Al-Sai and Laith Mohammad Abualigah, “Big Data and E-government: A review,” 2017 8th International Conference on Information Technology (ICIT)
- Kwon Yeong-il (Victor) (Head of National Big Data Center(NBDC)), Big data and its Application in Korea, 2017
- http://www.sthda.com/english/wiki/text-mining-and-word-cloud-fundamentals-in-r-5-simple-steps-you-should-know

Reference

- KOICA website: www.koica.go.kr
- Korean Government, Leading_the_world_e-Government (brochure)
- Rogers, Everett M. 2000. Informatization, globalization, and privatization in the new millennium. The Asian Journal of Communication Volume 10 Number 2. pp. 71-92.
- WEF(World Economic Forum) Report
- The presentation by Dr. Jim Spohrer (IBM) © IBM UPWard 2016)
- Kyeong Seok HAN and Seok Yong Yun, A Study on Establishing Competitive Advantage Strategies based on Patent Data Investigation using
- TF-IDF(Term Frequency - Inverse Document Frequency) and Network Analysis, GLOBAL CONFERENCE ON BUSINESS MANAGEMENT AND ECONOMICS 2018)
- https://www.privacyshield.gov/article?id=Korea-Cyber-Security

Reference

- Nora, Simon; Minc, Alain (1981). The computerization of society : a report to the President of France0-262-64020-1 (2. pr ed.). Cambridge, Mass.: The MIT Press. p. 208. ISBN 0262640201.
- Kim, Sangmoon. 2004. Social Informatization: Its Measurement, Causes and Consequences. Paper presented at the annual meeting of the American Sociological Association, Hilton San Francisco & Renaissance Parc 55 Hotel, San Francisco, CA, August 14, 2004 [Online PDF]. Available at <http://www.allacademic.com/meta/p110605_index.html. Accessed 19 July 2008
- https://www.itgovernance.co.uk/what-is-cybersecurity

References

- Gerard Salton(1988), "Term-weighting approaches in automatic text retrieval", Information Processing & Management Vol.24 Issue 5
- Japanese Ministry of General Affairs(2012), "Information Communication Technology White Paper in 2012"
- Jun Sunghae(2013), "A Big Data Learning for Patent Analysis", Journal of Korean Institute of Intelligent System Vol23 Issue 5
- KIPO(Korean Intellectual Property Office) (2018), http://www.kipo.go.kr
- KIPRIS(Korean Intellectual Property Rights Information Service) (2018), http://www.kipris.or.kr
- Luhn, Hans Peter (1957). "A Statistical Approach to Mechanized Encoding and Searching of Literary Information" (PDF). IBM Journal of research and development. IBM. 1 (4): 315
- Max H. Boisot, Ian C. MacMillan, Kyeong Seok Han (2008), "Explorations in Information Space", Oxford
- The Economist(2017), https://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource
- WIPO(2017), "Guide to the International Patent Classification"
- World Intellectual Property Organization, http://www.wipo.int

THANK YOU!

Q & A