The Issues of Big Data Network and AI Initiatives

February 16th, 2020

Kwon Yeong-il (Victor), Ph.D.
Professor of Hoseo University
Contents

1. Digital Transformation
2. Data Capital & Data Economy
3. Big data use cases in Korea
4. AI (Artificial Intelligence) Initiative
Digital Transformation

Innovative Technologies of 4th Industrial Revolution

IoT  Big Data  Chatbot
Hyper Connectivity  Block Chain  Autonomous Robotics
4th Industrial Revolution  Machine Learning
Deep Learning  Autonomous Driving  Robo Advisor
Smary Factory  3D Printing  Cloud Computing
Digital Transformation
Digital Transformation

New Paradigm of Manufacturing Industry

The Industry 4.0 based on the Digital Transformation

- Gearing the future
- Human as a machine
- Standardization, Economy of Scope

- Creating the future
- Machine as a human
- Uniqueness and Flexibility

A true platform doesn’t be built just by making a company bigger, rather by a thought helping others do better.
Digital Transformation
Beginning of New Digital Economy

Technology Convergence
- Robot
- AI
- IoT
- Big Data
- Cloud Computing
- Unmanned Car
- Block Chain
- 3D Printing
- Genome
- Neural Technology

Digital Transformation
- Personalization (B to P)
- Customized Product
- Cyber Physical System
- Digital Twin
- Industry 4.0
- Digital Innovation
- Shared Economy
- On-demand Economy
- O2O Economy
- Platform Revolution
Use of an ecosystem of digital technologies…

• Enables a new wave of digital transformation that:
  – Builds on digitisation and datafication
  – Is more than the sum of its parts
  – And includes technologies such as:
    • Big data
    • Cloud computing
    • Internet of Things
    • Robotics
    • 3D printing
    • Artificial Intelligence
    • Distributed ledgers
    • …
I don’t need a car, I need mobility.

With drones, I can get deliveries anywhere.

I can afford this house, by renting it out.

... is a game-changer providing new opportunities and enabling new business models.
... is disrupting the markets

Losers

Winners

[Images of stores with "STORE CLOSING" signs, a "Blockbuster Video" sign, a "Borders" sign, and "Goodbye Old Friend" sign on the left.

Images of "Amazon", "Netflix", and "Instagram" logos on the right.]
… also raising digital privacy and security issues
Data: A fundamental driver
Quantum-jump computing power
Data as a core driver of disruptive innovation

- The use of big data promises to significantly improve products, processes, organizational methods and markets, a phenomenon referred to as data-driven innovation.
Contents

1. Digital Transformation
2. Data Capital & Data Economy
3. Big Data Center & Use Cases in Korea
4. AI (Artificial Intelligence) Initiative
Data Capital & Data Economy

The Change of Data Value

Resource, Data refinement …  → New Capital of Digital Economy

The oil of the 21st century (Gartner, 2012)

The rise of 4th industrial revolution

The Rise of Data Capital

Data is New Capital of Global Economy
(Deloitte, 2013)

"Big data is the capital in the 4th revolution"

"Monetary capital is the key factor in the manufacturing industry, however, Big data is the most valuable capital" (Vice President, Roland verger, march, 2017)
Data Capital & Data Economy

Data is now a form of capital, on the same level as financial capital in terms of generating new digital products and innovative services.

Exponential Organizations (Unicorn)
Marginal Cost Zero, Winer-Takes-All (WTA)

< 3 factors of production >

Cyber (Internet)
Data Scientist
Data Capital

Land (Factory)
Labor (Skilled Worker)
Monetary Capital

Data Economy

vs

Industrial Economy
The Potential of Data Capital

- **Societal benefits** in many areas such as health, environment, agriculture, mobility, research, and society’s progress
- **Economic growth** in many business for competitiveness, innovation, job creation

*Bringing it all together!*

Combining government, industry and scientific data

Innovation & growth + solutions to societal challenges
Data Resources for building a Data Economy

1) Personal data - GDPR
   - Anonymized personal data: treated like non-personal data
   - Data protection logic
   - Free flow of personal data

2) Government data – PSI (Public Sector Information) & Open data
   - Avoid discrimination between re-users
   - Address re-use applications within a time limit
   - Limit use of exclusive arrangements
   - Limit charges (marginal cost of reproduction)

3) Research data – Open Science
   - avoid discrimination between re-users
   - Address re-use applications within a time limit
   - Limit use of exclusive arrangements

4) Industry-held data
   - Focus on non-personal, machine-generated data (ex, IoT data)
   - Contracts are main vehicles to share and re-use
   - Data silos innovation hampered
Data related with 4th industrial revolution

1) Type of Big Data
- Bio tech: genome data
- Medical: patient record, sensor data, image, video data
- Manufacture: MES data, IoT-device enabled data, CPS data
- Transportation: traffic data, DTG data
- Security & Safety: CCTV, 112 voice data, accident data
- Finance: credit/debit data, stock trading
- Energy: Electric power, smart sensor data
- Distribution: logistics data by RFID
- Administration: government-owned data
- Welfare: pension data
- Agriculture: IoT smart farm data

2) Sour of Big data
- Anonymized personal data: treated like non-personal data
- Sensor, CCTV
- Internet of Things (IoT)
- Wearable device: CGM (Continuous Glucose Monitoring)
- Monitoring tool: EMS, BEMS, ...
- Location: GPS
- Traffic: VDS (Vehicle Detection System), AVI (Automatic Vehicle Identification) system, TCS (Toll Collection System), Hi-Pass system
Ecosystem of Open Data

8 dimensions considered essential for an open data initiative that builds a sustainable open data ecosystem

Source: Open Data Readiness Assessment: Malaysia
Global LOD (Linked Open Data) Cloud Diagram

Web of Internet ⇒ Web of Data!

http://lod-cloud.net/
SMEs in a digital economy

Big Data Analysis for SMEs is challenging!

Mostly Uses external data rather than internal data!

Big data Solution Matching Project (No. of Company)

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Internal (37%)</th>
<th>External (63%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT System</td>
<td>3 17%</td>
<td>2 20%</td>
</tr>
<tr>
<td>Sensor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SNS</td>
<td></td>
<td>63%</td>
</tr>
</tbody>
</table>

1. **SNS Data**
   - Online Data (Blog, Twitter, News, Cafe, Community)
   - Thesis
     - Main Target Customer
     - Purchase Purpose
     - Product and Brand Image
     - PR Channel

2. **Sensor Data**
   - Wi-Fi Signal of Smart Phone
   - Identifier, Customer location

3. **IT System Data**
   - Transactional Data (ERP, POP, etc.)
   - Workflow Time, Facility Capacity, Production Data, Idle time etc.
SMEs in a digital economy

Big Data Issues of SMEs

Roadblocks to Big Data

- Not much of data accessible while still wondering about the tangible impact of big data.
- Especially, SMEs think that big data is NOT quite relevant to what they do for their business.

N=560 with multiple answers

- Lack of Data: 289
- Lack of Relevance to What the Agency/Company does: 176
- Lack of Interest on the part of CEO/CIO: 122
- Lack of Confidence in Big Data Impact: 112
- Lack of Data Scientist: 92
- Lack of Understanding About Big Data: 78
- Lack of Work Related to Big Data: 72
SMEs in a digital economy

Explosion of Unicorn Companies ($1 Billion) in Global

Fortune 500 companies take 20 years, but, Unicorn companies takes 4.4 years in average

The Increasingly Crowded Unicorn Club
Private Unicorns since 2011
Contents

1. Digital Transformation
2. Data Capital & Data Economy
3. Big data use cases in Korea
4. AI (Artificial Intelligence) Initiative
Case 1: Bus line Routing Decision Support

KT + Seoul City Government

System Architecture

KT Platform

Hi’Seoul
Seoul Pedestrian Flow

PEDERSTRIAN ESTIMATION
NIGHT BUS LINE, STATION
ROAD

Bus line Routing Decision Support
Case 2: Using big data to monitor epidemic transmission

**Cause of Animal Epidemic Transmission**

- First Outbreak in Gimbê
- 2/4–2/8 Car Visit 2
- 2/16 Outbreak
- 2/13 Car Visit 3
- 2/13 Outbreak
- First Outbreak in Suncheon
- 2/27 Outbreak

**Spread by vehicles visiting the infected farm**

**Cause of Human Epidemic Transmission**

- Spread by people visiting the infected area

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

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

source: European Centre for Disease Prevention and Control
Case 2: Infectious Diseases Monitoring

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
Case 3 : Traffic Accident Prediction

- Collect and analyze various internal and external data including traffic accident data
- Develop traffic accident risk prediction system by applying deep learning technology with selected traffic accident variables
- Inform through TBN Korean Traffic Broadcasting and navigation contents

---

Data Collection/Process   Pretreatment and Formation statistics   Traffic accident risk prediction analysis system   Storing data mart & opening data   Web visualization & utilization

<table>
<thead>
<tr>
<th>Bigdata Platform</th>
<th>Established Traffic Accident Prediction System Based Bigdata</th>
<th>Risk Prediction Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Accident</td>
<td>Development of traffic accident risk prediction algorithm</td>
<td>Operate the Web dashboard service</td>
</tr>
<tr>
<td>City traffic</td>
<td>Application of technology to predict traffic accident risk</td>
<td>Provide forecasting data</td>
</tr>
<tr>
<td>The Public 3.0</td>
<td>Ensemble modeling</td>
<td>Provide forecasting data</td>
</tr>
<tr>
<td></td>
<td>Deep Learning Algorithm</td>
<td>Expansion of target service area</td>
</tr>
<tr>
<td></td>
<td>Analysis processing technology</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GPU Parallel processing</td>
<td></td>
</tr>
</tbody>
</table>
Case 3: Traffic Accident Prediction

Selection of Traffic Accident Variable and Design Deep Learning Algorithm

Selection of traffic accident variables

Use existing traffic accident causation pattern analysis

- [Bayesian network]

Local environment variables
- Road type and shape, road width, road line shape, intersection shape, etc. (13 variables)

Accident pattern variable
- Time of accident, vehicle use, driver profile, pre-accident behavior, etc. (20 variables)

External variable
- Traffic, weather, interdiction equipment, demographics, traffic culture index by region, etc. (10 variables)

Deep Learning Analysis Algorithm

- Business viewpoint
- Variable definition
- Data perspective variable selection (except missing, outlier)
- Final selection of variables of utilized analysis perspective

Calculation of accident risk prediction probability at 4,500 points in 6 regions

Local environment variable
- Intersection type
- Road type
- Traffic volume
- Weather condition

Accident pattern variable
- Local environment variable
- Accident pattern variable

External variable
- Input Layer

FM Model
- FM Layer
- Addition
- Inner Product
- Sigmoid

DNN Model
- Embedded DNN Model
- Hidden Layer
- RELU

Output Layer
- Embedded Layer
- Input Layer

Intersection type
Local environment variable
Accident pattern variable
External variable
Case 4: Jeju Big Data Platform Project

Organize local industrial ecosystem using data service

Activate local data industry

Establish local data sharing system

Collect practical public data and extend the range of local data services

Foster Private sector success capability

Develop local public life service

Establish customized data service considering local environment

Share data management strategies & techniques of active private companies
Case 4: Jeju Big Data Platform Project

Open Data

Move + Search
+ Bus/Station Wifi
+ Card Usage
= Local Full Routing

KaKaoT
Korea’s No 2 Navigation Provider

Public Wifi Usage Data
Case 4: Jeju Big Data Platform Project

Public places + Bus and bus platform Wifi access information
• Bus transportation cards information
• BC card consumption pattern information

kakao
• Movement data of Kakao(bus, automobile, taxi)
• Search data of Kakao

Public institutions
• Jeju Agricultural Technology Institute
  + Meteorological Administration: Jeju Detailed Weather Information
• Information of related organizations such as Jeju Tourism Corporation

"Full routing information mashup of floating population according to Jeju situation"
Case 4: Jeju Big Data Platform Project

Private institution

Jeju Provincial Government

Kakao Co., Ltd.

Community-based public-private convergence data service

Service platform

API Gateway

Data Sharing

Data Standardization

Data Visualization

Configure test services

Data Management

API Management

Development Support

Data Connection

Public Data Portal

Data Application Service

Data Visualization

Data Sharing

Data Sharing

Data Sharing

Data Sharing
Contents

1. Digital Transformation
2. Data Capital & Data Economy
3. Big data use cases in Korea
4. AI (Artificial Intelligence) Initiative
Paradigm Shift to AI

AI is already Here!

- AlphaGo Game (March 9th-15th, 2016, Seoul, Korea)
Machine Learning, Data Science, and Statistics
Techniques of AI (Artificial Intelligence)

- **Learning**
  - Knowledge base
  - Learning model (Machine Learning)
  - Natural language processing

- **Inference**
  - Inference engine
  - Expert system
  - Pattern recognition & understanding system

- **Recognition**
  - Character, Speech, Image processing

A process of acquiring or modifying knowledge, behaviors, skills, values, or preferences.

A conclusion reached on the basis of evidence and reasoning.

A capability of seeing, listening and talking.
AI Landscape 2019 in Global

Enterprise AI Companies
PRESENTED BY TOPBOTS @mikequindazzi
Big Data Network and AI

Widening AI Gap

- Machine Learning (AI)
- DATA
- Service Evolution
- Service
- User

AI evolution based on data network effects
Wider technology gap

AI evolution without data network effects
Strategic Plan for Innovative Growth

Key of 4th Industrial Revolution - AI, BigData, IoT

- AI: Big Data based Learning
- Big Data: Data Collection/Distribution/Analysis
- IoT: Data creation/ Real time Monitoring
  - 5G and 6G Network
AI Learning Knowledge base Infra (2017-2021)

AI Data Infrastructure

AI Learning Data
- Refinement
- Modeming
- Analytics

Tagging
- 1. Knowledge base: Medical, Patient, Legal...
- 2. Computing: Power
- 3. Korean ImageNet

Data Dictionary

Business Area
- AI Startup
- Institute
- SW Ventures
- Hospital
- Data Companies

AI Project
- ExoBrain
- IBM Watson
- Buno
- ICT Projects
- R&D Project

Phase 1 (2017)
- Medical
- Patient
- Law

Phase 2 (2018)
- Finance
- Education

Phase 3 (2019-2021)
- Agriculture
- Energy
- Logistics
- Manufacture
- Transportation

Diffusion of AI Knowledge

AI Knowledge Development
Big Data Network Initiative

Data as Core Asset for Innovative Growth

We will build a Korea Data eXpress (KDX) for the Data Economy Era, just like it built the highway road in the industrial age.

"From Internet savvy to data savvy, to a safely handled data country – this is what we want to be"

The Korean Government will fully support the data industry to bring a new vitality to our economy.

President Moon Jae-in’s Declaration on regulatory reform for data economy vitalization (August, 2018)
Change of Key Economic Components by age

Core of the Innovation – Big Data eXpress

[1970's] Industrialization Age
(Road) Expressway (SOC)

[2000s] Information Age Information
Superhighway (Network)

[2020] Intelligent Age
Korea Data
eXpress(Data)
Strategy for Big Data eXpress (1)

Build & Open up High-value & High-demand Data

01 Open all public data

02 Generate and construct data for AI learning

03 Standardize public data and improve quality
Strategy for Big Data eXpress (2)

01 Expand data cooperation network
02 Facilitate private cloud use
03 Establish foundation for data trading
Key Success Factors (KSF) of AI

✓ 2 Key Factor for AI Success
  • Time To Market!
  • AI is a CEO Agenda!

✓ 4 Success conditions for AI

Big Data
Data is the New Capital
50 Zetabytes created in 2020

New Algorithms
Massively parallel delivering Superman Accuracy

Super Computing
Massively Parallel Architecture Driving Performance

Data Expert
Analyzing Big data and Visualization

![Image](image.png)
Key Driver of Industry 4.0

Industry 4.0: Key Driver [Skill, Policy, Culture]

- **Skill**
  - Reskill / Upskill / New skill
  - Syllabus & Curriculum revision

- **Policy**
  - Incentive, Data security
  - Foreign labour, Digitalisation, ICT infrastructure

- **Culture**
  - New paradigm shift in corporate culture
  - Change process
Appendix: Requirements of Data Architect
**ABL (Actual-task Based Learning) Program in Hoseo University with MoI and MoICT**

1. ABL Curriculum
   - Master/Ph.D. with Enterprises/Governments

2. Peer ABL

3. Field ABL

4. ABL Day

5. ABL Forum

**Global Partnership**
- ASEAN, Industrial HRD in Graduate School

**R&D Project**
- Domestic/Overseas University/Institute

**4th IR Skill-up**
- AI/Big data, IoT, Platform business, Smart Factory

**Global Tech-Business and HRD 4.0 Model at Hoseo Univ.**

**Smart Product/Service**
- High-tech business
  - Big data/Al/Smart factory, etc.

**Cost Saving, Process Innovation**
- Digital transformation

**Recourse Sourcing**
- Funding, Skilled employees

**Global Tech-business**
- Business cooperation
ABL Education Model in Hoseo University

Field ABL (Actual-task Based Learning)

< Training of Smart factory (6 Weeks)>

Global Tech-business Forum 2019
(September 28th)

SMEs (Industry)

ABL (Hoseo University)

4IR Projects (Government)
Terima Kasih!

Dr. Kwon Yeong-il (Victor)

kyi@hoseo.edu