KCS' big data and AI strategy
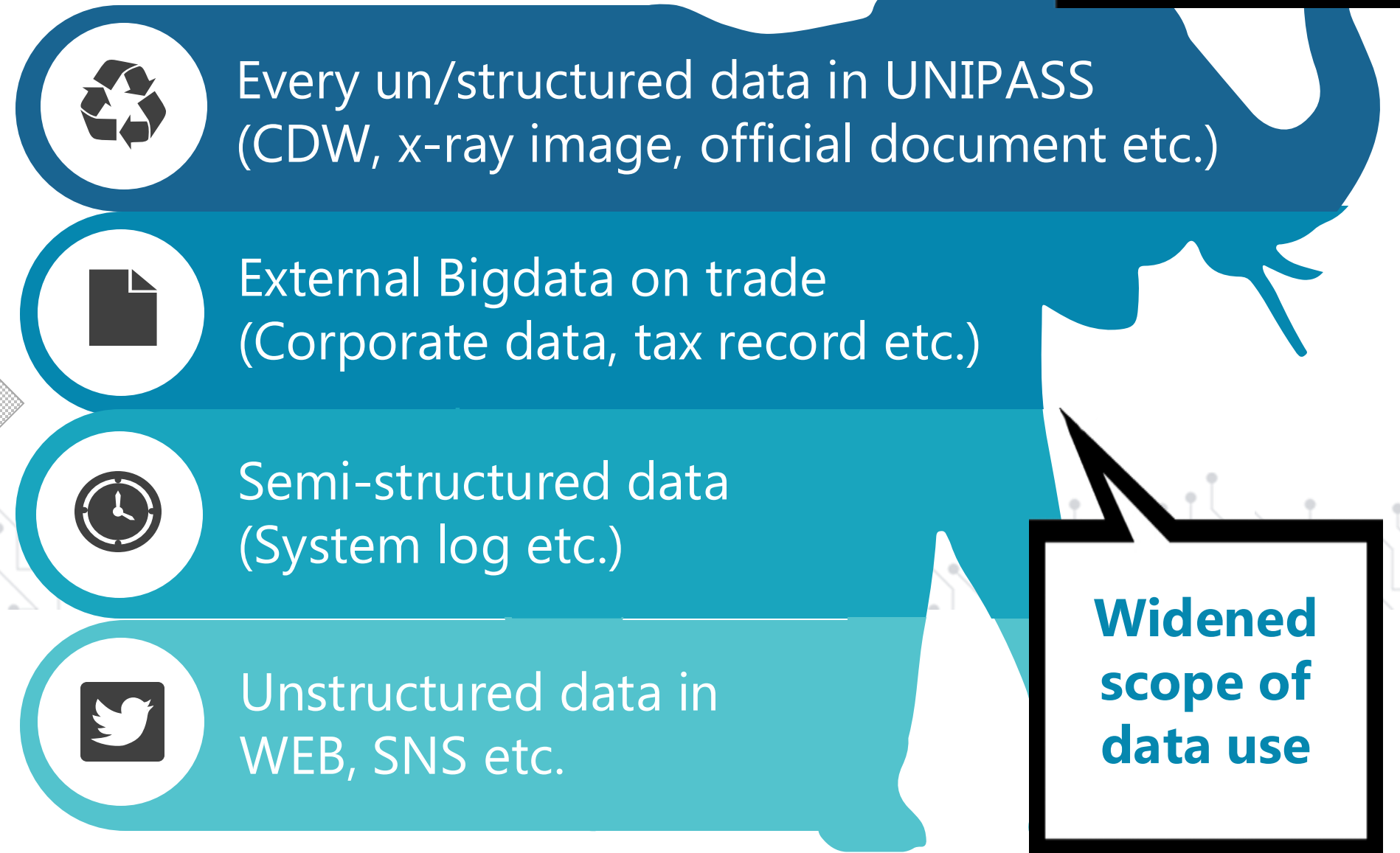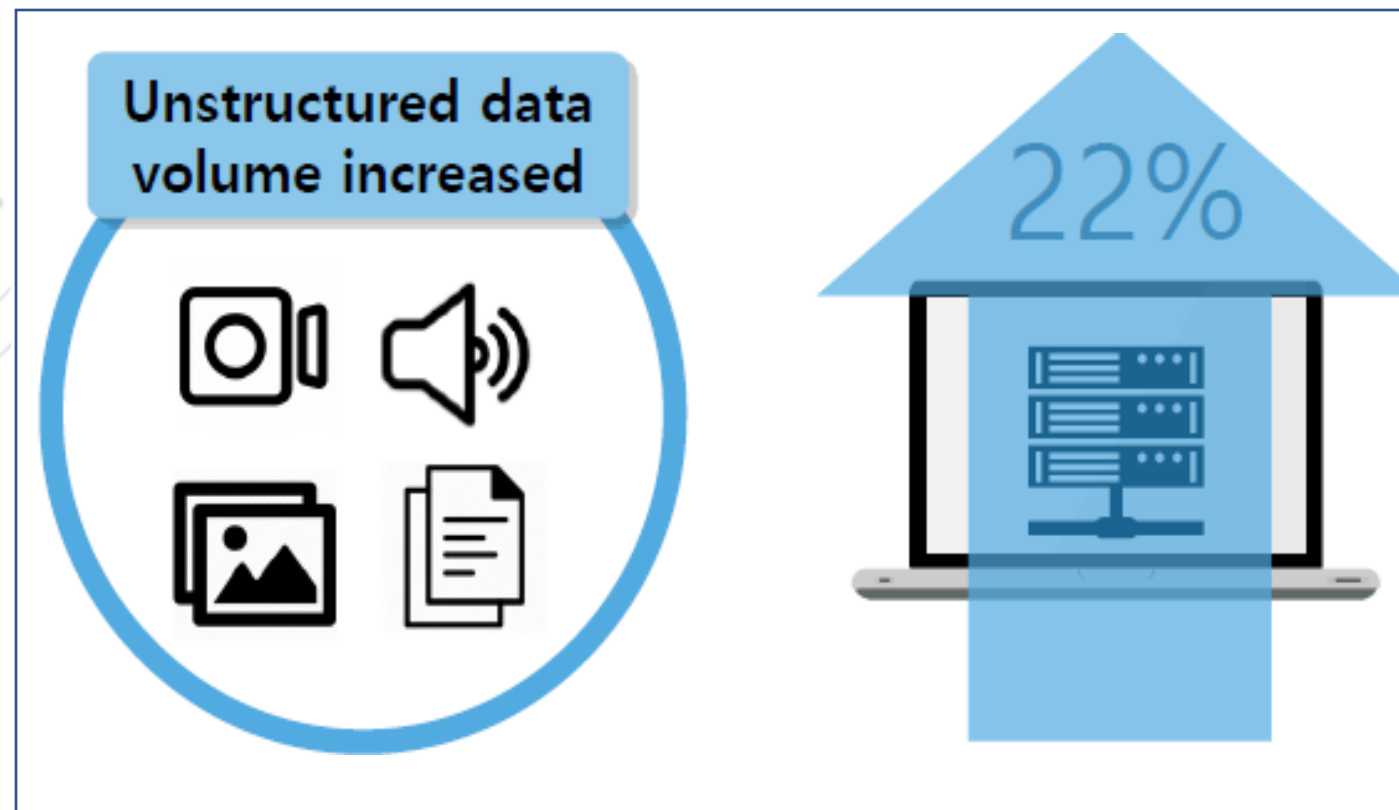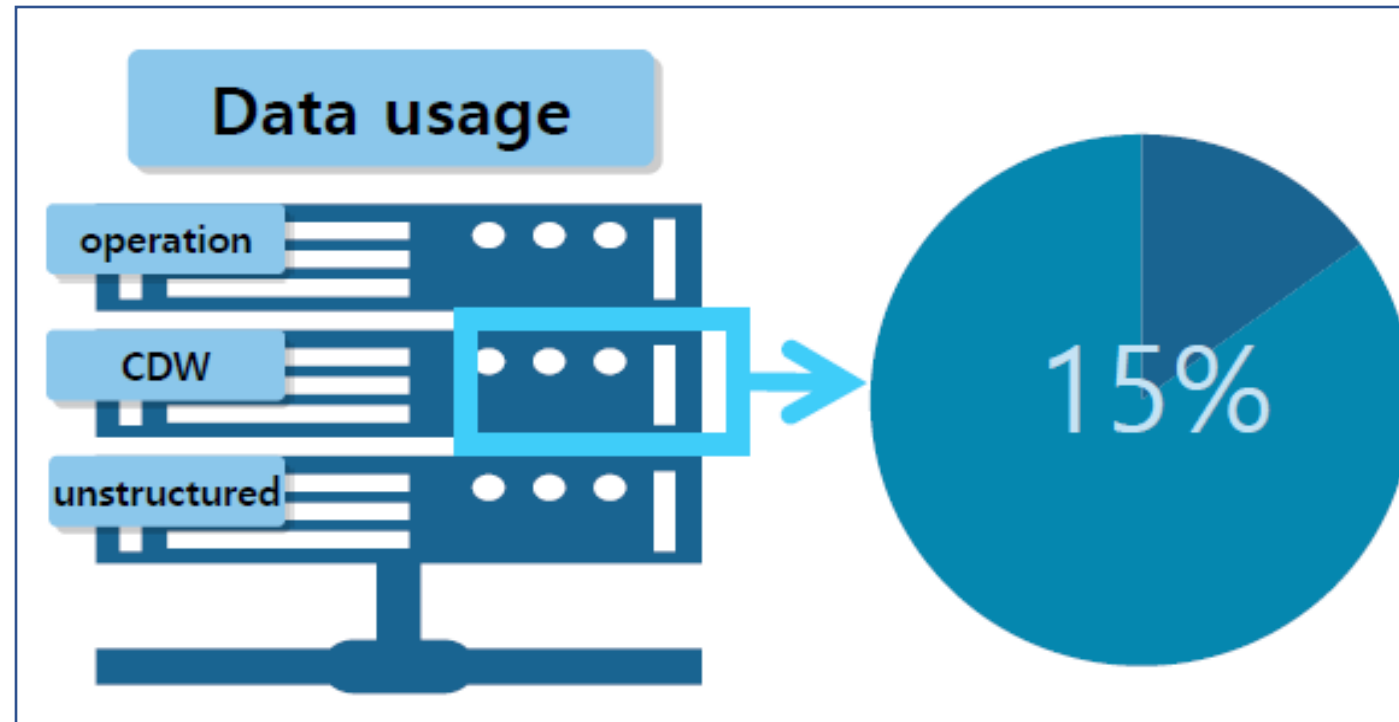# in controlling risks on border operations

관세청

Miryang Kim, 24.10.23
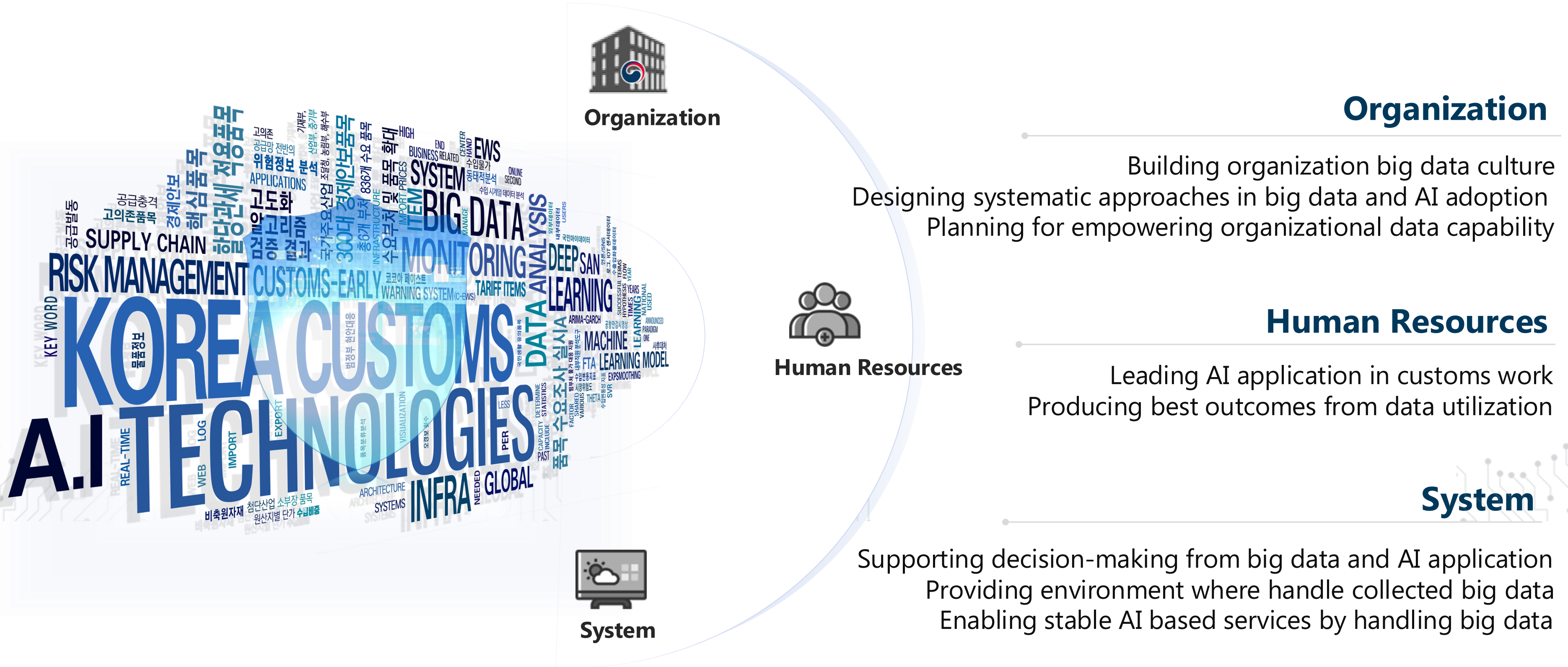
Paradigm Shift to Big Data

✓ Traditional data analysis reveals the limitation of its effects, utilized data and skills.

Data usage
- operation
- CDW
- unstructured

15%

Unstructured data volume increased

22%

Various advanced analysis

- Every un/structured data in UNIPASS (CDW, x-ray image, official document etc.)
- External Bigdata on trade (Corporate data, tax record etc.)
- Semi-structured data (System log etc.)
- Unstructured data in WEB, SNS etc.

Widened scope of data use

Key Points of Big Data Strategy

✓ Korean customs focuses on 3 aspects to leverage big data and adopt AI technology.

Organization

Human Resources

System

Organization

Building organization big data culture
Designing systematic approaches in big data and AI adoption
Planning for empowering organizational data capability

Human Resources

Leading AI application in customs work
Producing best outcomes from data utilization

System

Supporting decision-making from big data and AI application
Providing environment where handle collected big data
Enabling stable AI based services by handling big data

✓ Designated organizations should drive the direction of technology adoption.

Big Data Analysis Division(2020.7.14)

Tasks

- **Maintaining a big data platform**
  - Providing services by applying AI and data skills
  - Operating platform to guarantee sufficient functions for AI models and analyze big data
- **Designing ICT expert training programs**
  - Organizing events to deliver advanced technology trend
- **Adopting Generative AI**
  - Building a master plan for AI adoption and performing PoC

Performance system

- **Headquarters**
  - Leading technology adoption, planning and supporting big data utilization
    Providing adequate environment
- **Branch customs office**
  - Developing cases of data utilization and finding areas requiring data-based performance
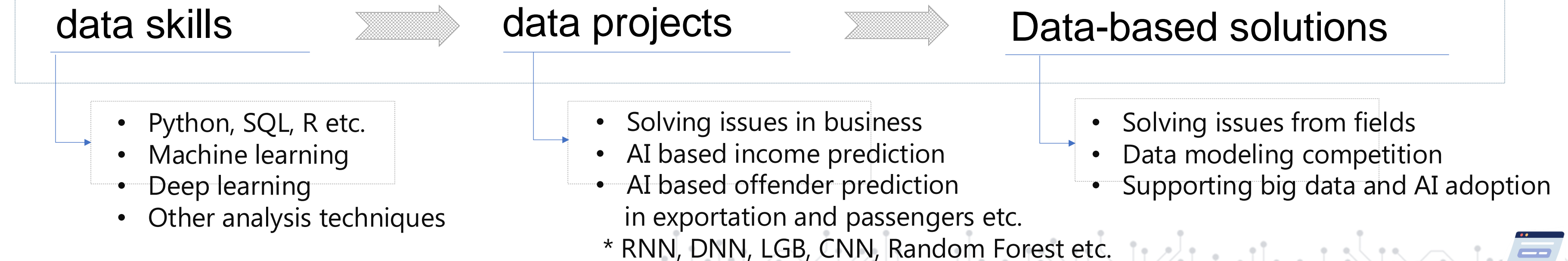
Human Resources

✓ Capability on technology and domain is the key aspect allowing us to achieve the desired results from AI application

ICT Expert Training Program(2017~)

- **145** Trainees
- **53** Data projects

Curriculum

data skills ➡ data projects ➡ Data-based solutions

- Python, SQL, R etc.
- Machine learning
- Deep learning
- Other analysis techniques

- Solving issues in business
- AI based income prediction
- AI based offender prediction in exportation and passengers etc.
  * RNN, DNN, LGB, CNN, Random Forest etc.

- Solving issues from fields
- Data modeling competition
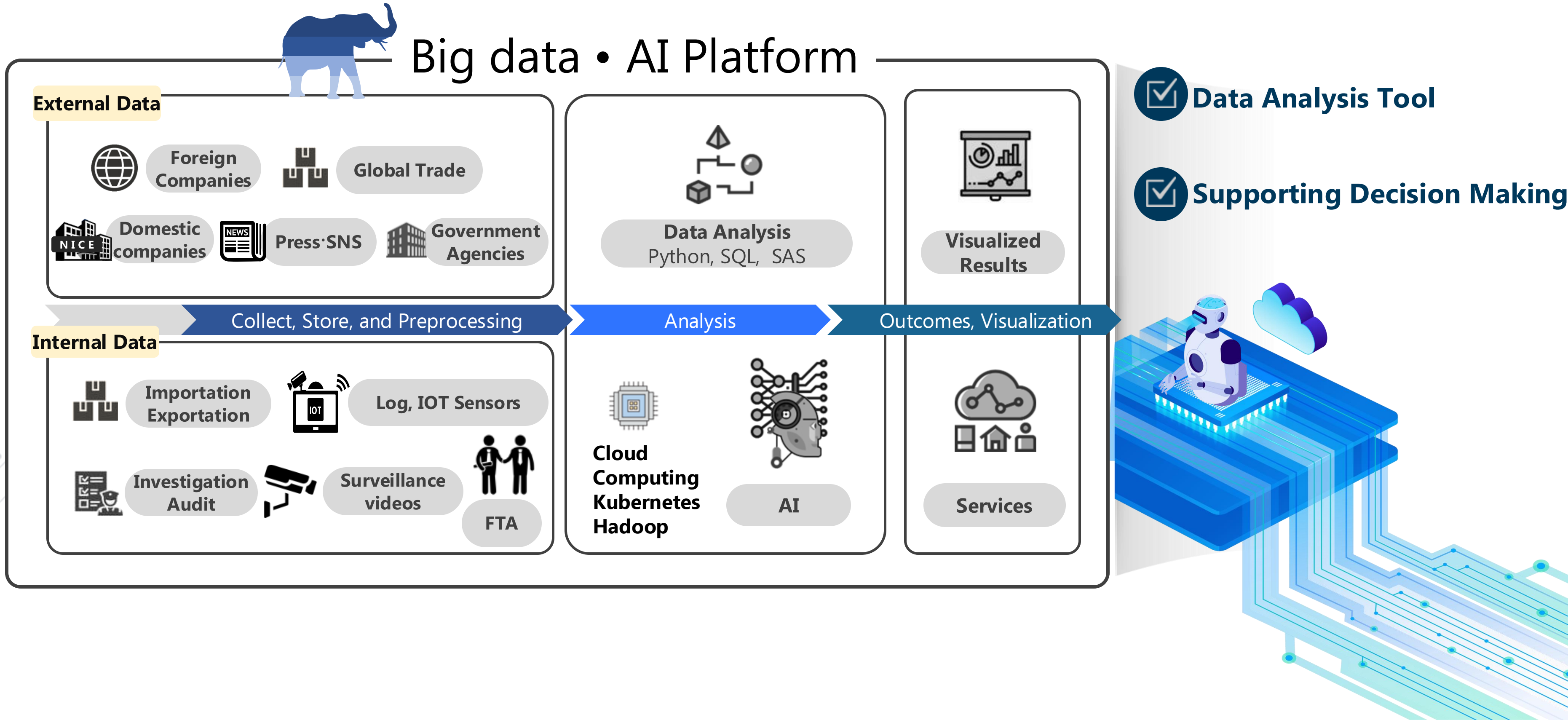- Supporting big data and AI adoption

Establishing big data culture

- Fostering organizational capability in data utilization
- Providing diverse lectures delivering insights on cutting-edge technology

Artificial Intelligence

Systems

✓ System environment allows handling big data and providing services based on AI technology

Big data • AI Platform

External Data

- Foreign Companies
- Global Trade
- Domestic companies
- NICE
- NEWS Press·SNS
- Government Agencies

Internal Data

- Importation Exportation
- IOT Log, IOT Sensors
- Investigation Audit
- Surveillance videos
- FTA

Collect, Store, and Preprocessing → Analysis → Outcomes, Visualization

Data Analysis
Python, SQL, SAS

Cloud Computing Kubernetes Hadoop

AI

Visualized Results

Services

☑ **Data Analysis Tool**

☑ **Supporting Decision Making**

AI Application in Korean Customs

✓ Korean Customs applies AI by using diverse data for risk management.

Data
- Structured data: data in relational database
- Unstructured data: texts, images, videos etc.

Technologies

Rule-based ⟹ ✚ Advanced Data Skills

- Statistical algorithms
- Machine learning
- Deep learning
- Natural language processing

- Classification
- Network Analysis
- Image detection
- Similarity analysis

Services
- Risk Management: Items, companies, and humans

- AI Technologies: Random forest, SVM, CNN, RNN, NLP, and others

- Data: Structured(Database: relational, graph)
300TB + Unstructured data(Images, videos, texts)

✓ AI models play the prominent roles in the risk management by controlling risk factors of importation

AI Customs Clearance Model

Finding risk levels of importation

- Machine Learning and Deep learning
- Generating group of high-risk importation for a certain period of time
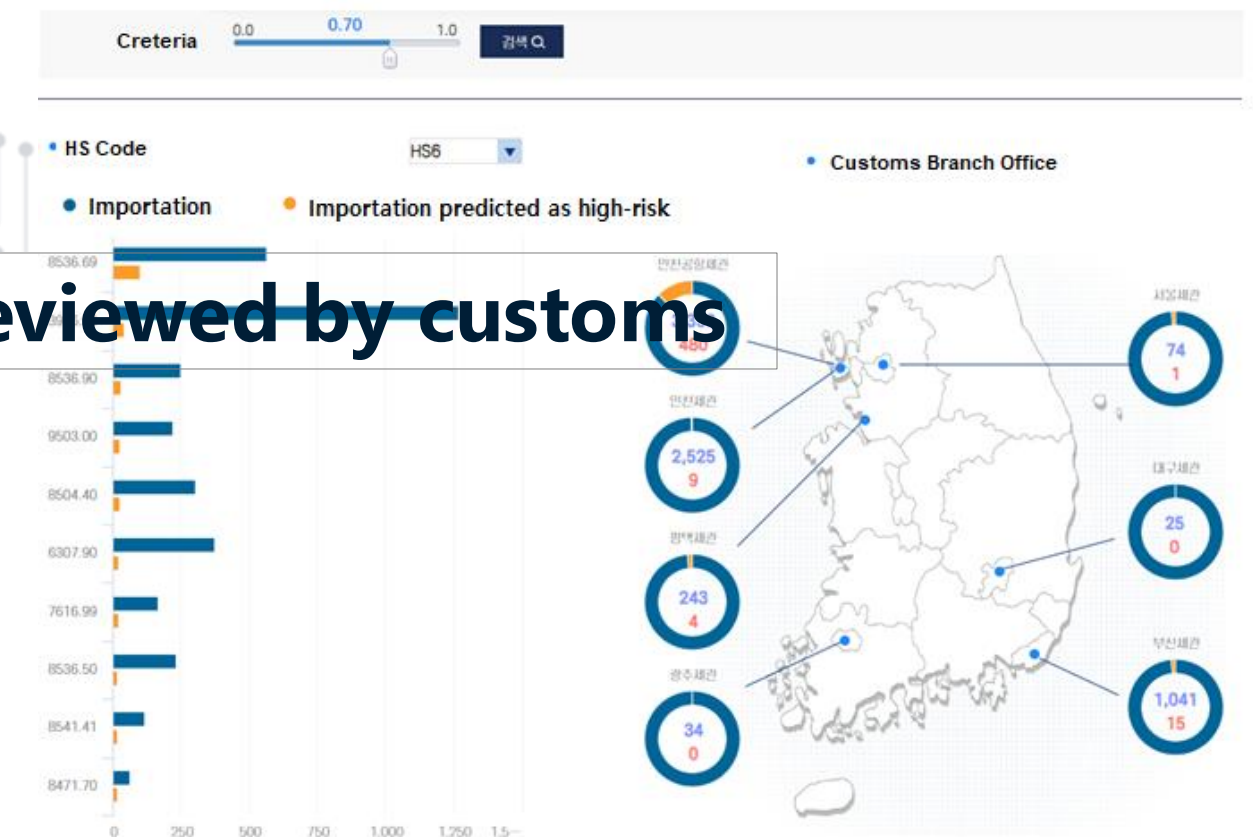- Measuring level of risks in importation by learning pattern of data
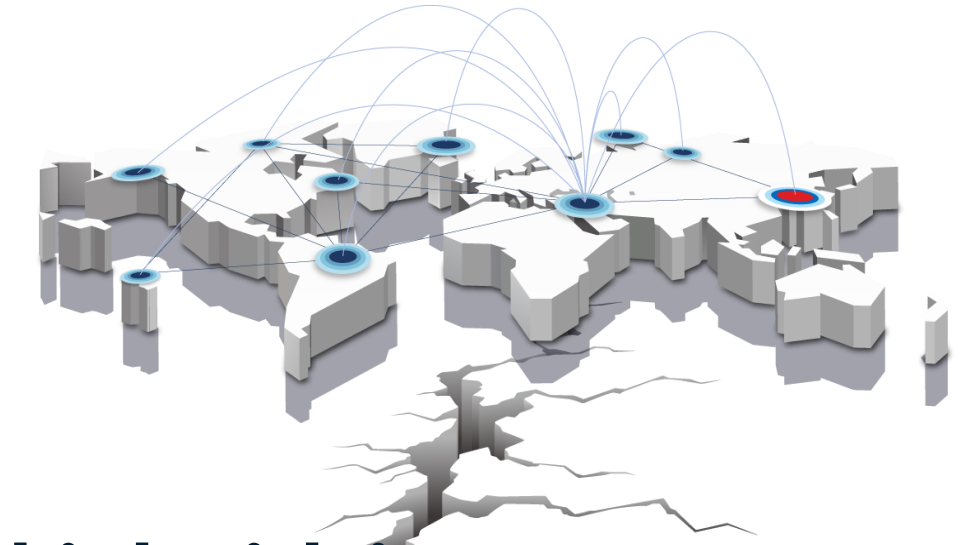
➡ **Focusing on high-risk items processing safe one through fast track**

AI Prediction of items with possible risk

Finding items should've reported as the object to be reviewed by customs

- Machine learning
- Predicting items that have the possibility of being falsely declared as general import

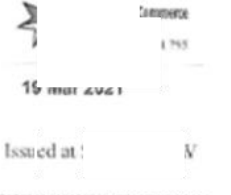* Clustering, FCNN, Random Forest, LightBGM, SVM, RNN etc.

✓ Diverse data and technologies are used in AI modelling.

AI Identifying Counterfeit Documents

Comparing Similarity of Images

- CNN-based models
- Detecting objects(stamps and signatures from images) from the given images
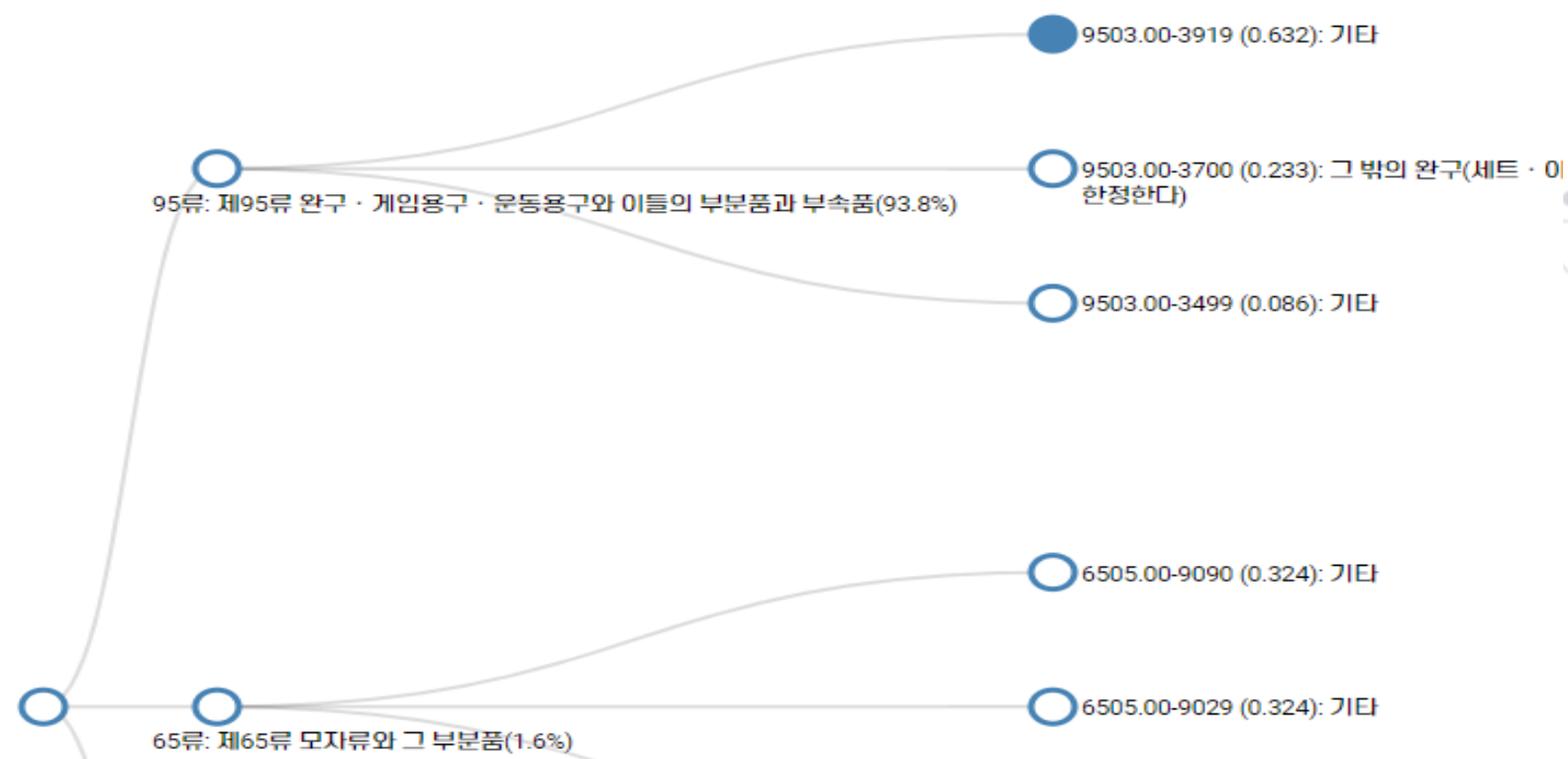- Comparing the similarity of extracted images to the registered ones

Identifying counterfit Stamps/Signatures

	Issuing Authority Stamps		Issuer's Signature	
Targets				
1 Original				
Results	60.05%	25.99%	49.17%	100%

9503.00-3919 (0.632): 기타

9503.00-3700 (0.233): 그 밖의 완구(세트ㆍ이 한정한다)

95류: 제95류 완구ㆍ게임용구ㆍ운동용구와 이들의 부분품과 부속품(93.8%)

9503.00-3499 (0.086): 기타

6505.00-9090 (0.324): 기타

65류: 제65류 모자류와 그 부분품(1.6%)

6505.00-9029 (0.324): 기타

AI HS(Commodity Classification Code) Prediction Model

Recommending possible Commodity Code

- Random forest, SVM, Logistic regression
- Suggesting possible HS codes using trained models based on the declared item names, descriptions, and their HS codes

Thank you

singgi@korea.kr