Contents

KISA KOREA INTERNET & SECURITY AGENCY

I

Cybersecurity
AI Dataset
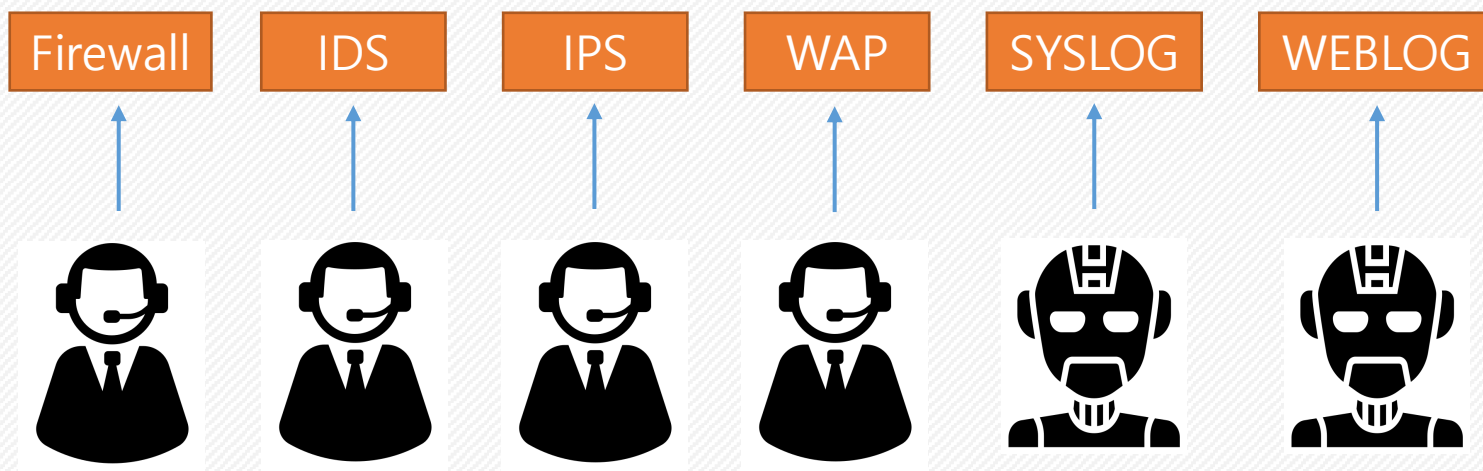# Introduction

01 AlphaGO(2016)

1. AI Dataset for Cybersecurity

02 Need for AI in Cybersecurity

➡ Expansion of Security data(Firewall, IDS, IPS , WAP, System Log, Web Log, etc.)

| Firewall | IDS | IPS | WAP | SYSLOG | WEBLOG |

03 **Require high-quality cybersecurity AI learning datasets**

Artificial intelligence will contribute to saving money and time by autonomously identifying or responding to potential cyber attacks (WEF, 2020)

☞ **The results of the AI model determine the quality of the learning dataset**

1. AI Dataset for Cybersecurity

04 Result of Cybersecurity AI Dataset(1/2)

Category		Dataset based Diagnostic Name of Malware	Dataset based Attribute Name of Malware	Dataset based Malware of resent Social Issues
Malware	Number Of Data	3 Hundred Million	1 Hundred Million	120,000
	File Type	More than 24 file types Include EXE	More than 30 file types Include PDF	More than 10 file types Include EXE,
	Labelling	More than 11,800 Family Dataset	More than 3,717 Similar/Variant Dataset	26 Keyword of recent Incidents

KISA KOREA INTERNET & SECURITY AGENCY

04 Result of Cybersecurity AI Dataset(2/2)

Category		Dataset based Detection of Cybersecurity Incident	Dataset based Cyber Attack Tactics	Dataset based Reenact of recent Cybersecurity Incident
Incident	Number Of Data	2 Hundred Million	1 Hundred Million	1.2 Hundred Million
	Environment	6 Heterogeneous Equipment Operating Environment (FW, IDS, IPS, WAF, etc)	Utilize an automated malware analysis platform	6 Heterogeneous Equipment Operating Environment (FW, IDS, IPS, WAF, etc)
	Labelling	More than 17 normal and aggressive acts	More than 230 types based on attack techniques	15 Cybersecurity Incident Scenarios

05 Verification of Cybersecurity AI Dataset

2021
사이버보안
AI 데이터셋
활용 우수 사례집

www.kisa.or.kr

과학기술정보통신부
Ministry of Science and ICT

KISA 한국인터넷진흥원
KOREA INTERNET & SECURITY AGENCY

Sharing 8 Best Practices

➡ Cooperate with Private/Public Cybersecurity Organizations

➡ Spam Filtering with Teleco,
Malware Detection with Game Publisher,
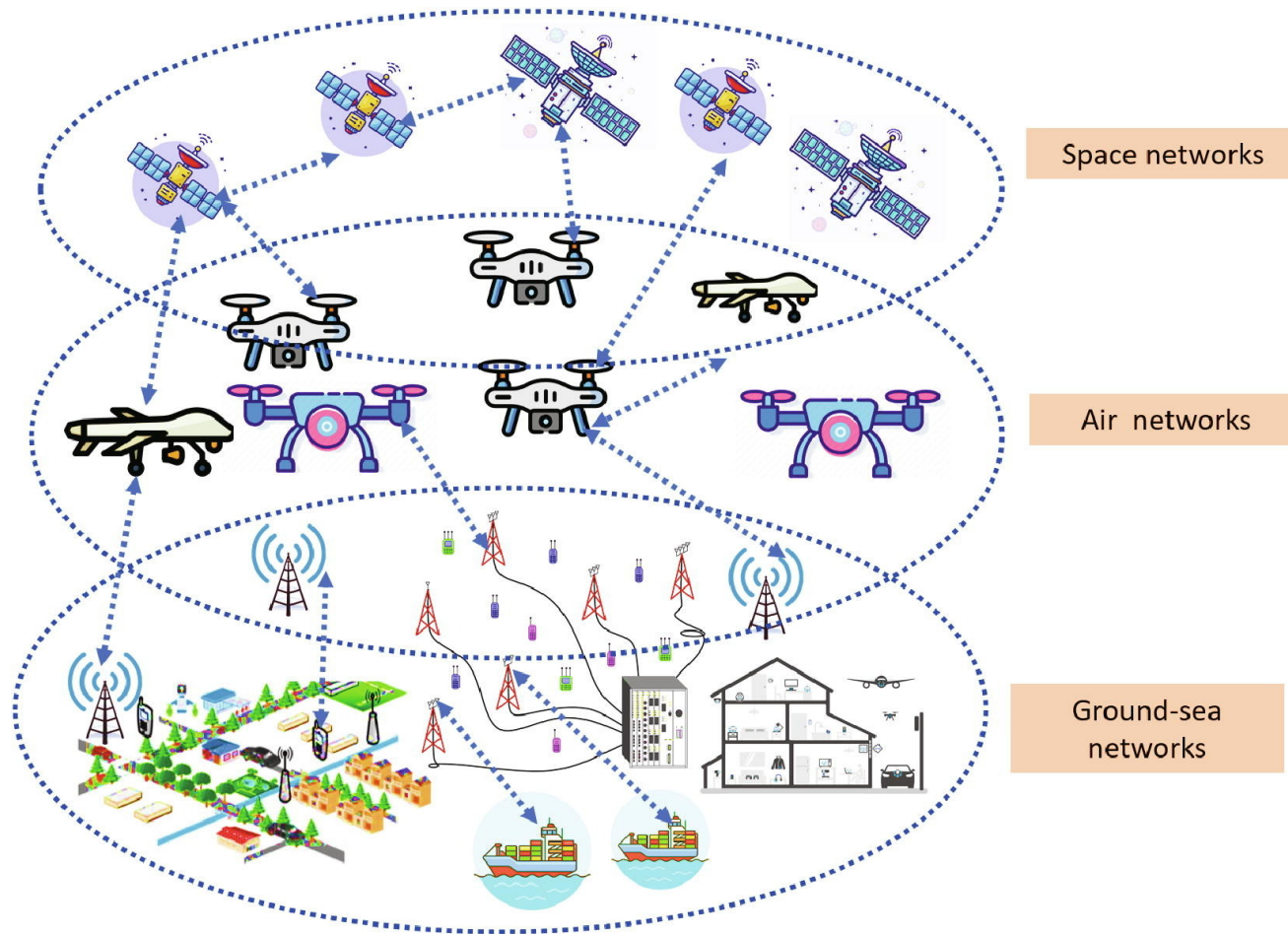Intrusion Detection with Monitoring Org.(CERT),
Etc.

☞ https://www.youtube.com/watch?v=nVijOJD3Efk

II

Cybersecurity
AI Dataset
In 6G

Shimaa A.Abdel, Hakeemab Hanan, H.Husseinb, HyungWonKim(2022)
Vision and research directions of 6G technologies and applications

01 Our Cybersecurity AI Dataset is good in South Korea Environment, But how about in other countries?

02 Device is vulnerable in the early stage of booting, Because still Worm Virus like WannaCry exist in Internet.

☞ Make standard procedure for device booting sequence, And standard format about Cybersecurity AI dataset.

KISA KOREA INTERNET & SECURITY AGENCY

Q/A

jmlee@kisa.or.kr