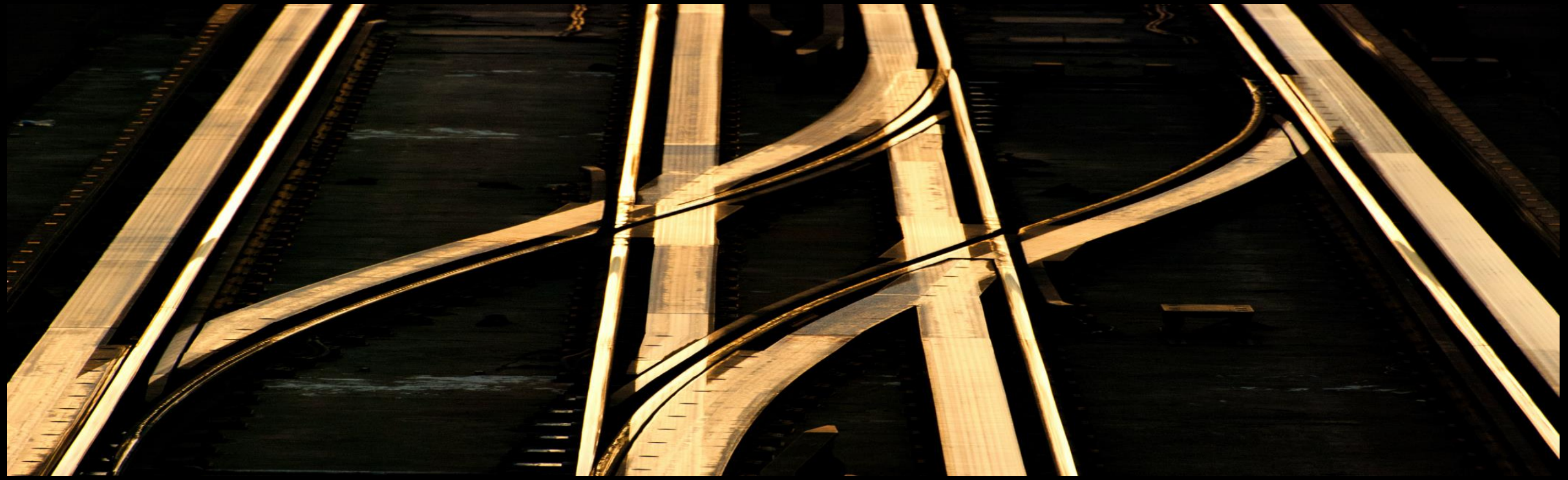


# AI and Security

Michele Bezzi, SAP Security Research
October 2019

PUBLIC

THE BEST RUN SAP

# AI & Security

**AI for  Security**:

- Size: …numbers ask for that…
- Complexity everywhere: It is not about the detection and response, only: AI for Secure Dev Lifecycle
- Humans are here to stay: explainable AI

**Security for AI**:

Attacks (and defense) to AI

© 2018 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

# SAP Security – Speeds and Feeds

## SAP Cyber Defense (statistics per month)

**Level 1 Outer Shell**

- **15,000,000+** malicious emails blocked
- **20,000,000+** threats blocked at the border of our network

**Level 2 Mid Layer**

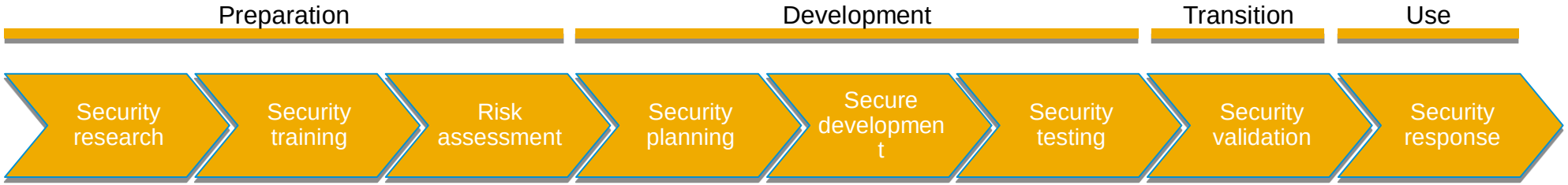
- **160,000+** malware detected and cleaned up by Anti Virus
- **200,000,000+** internet connections blocked (~2% of total)

**Level 3 Security Monitoring Center**

- **3.5+ TByte** log volume collected **per day**
- **30,000,000,000+** events collected in central system, thereof
  - **4,500+ Security events** correlated
  - **240+ Security incident** cases generated and defended

© 2018 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

# Open Source Management




**25,000+ SAP researchers and developers**

in over **130 countries**

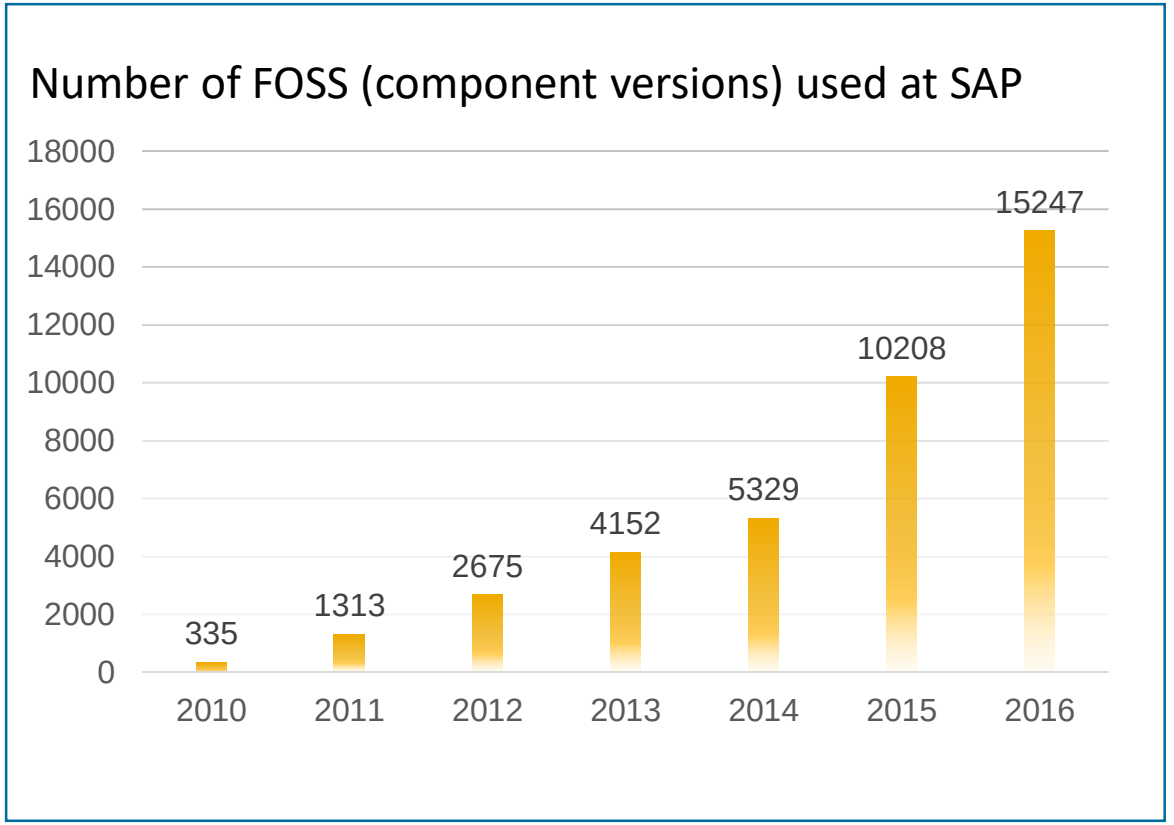
**100+ development locations** worldwide

**30+ acquired** companies in the last 10 years

**17,000+ SAP Partners** worldwide

**80%** of JAVA codebase is **Open Source** [2)]

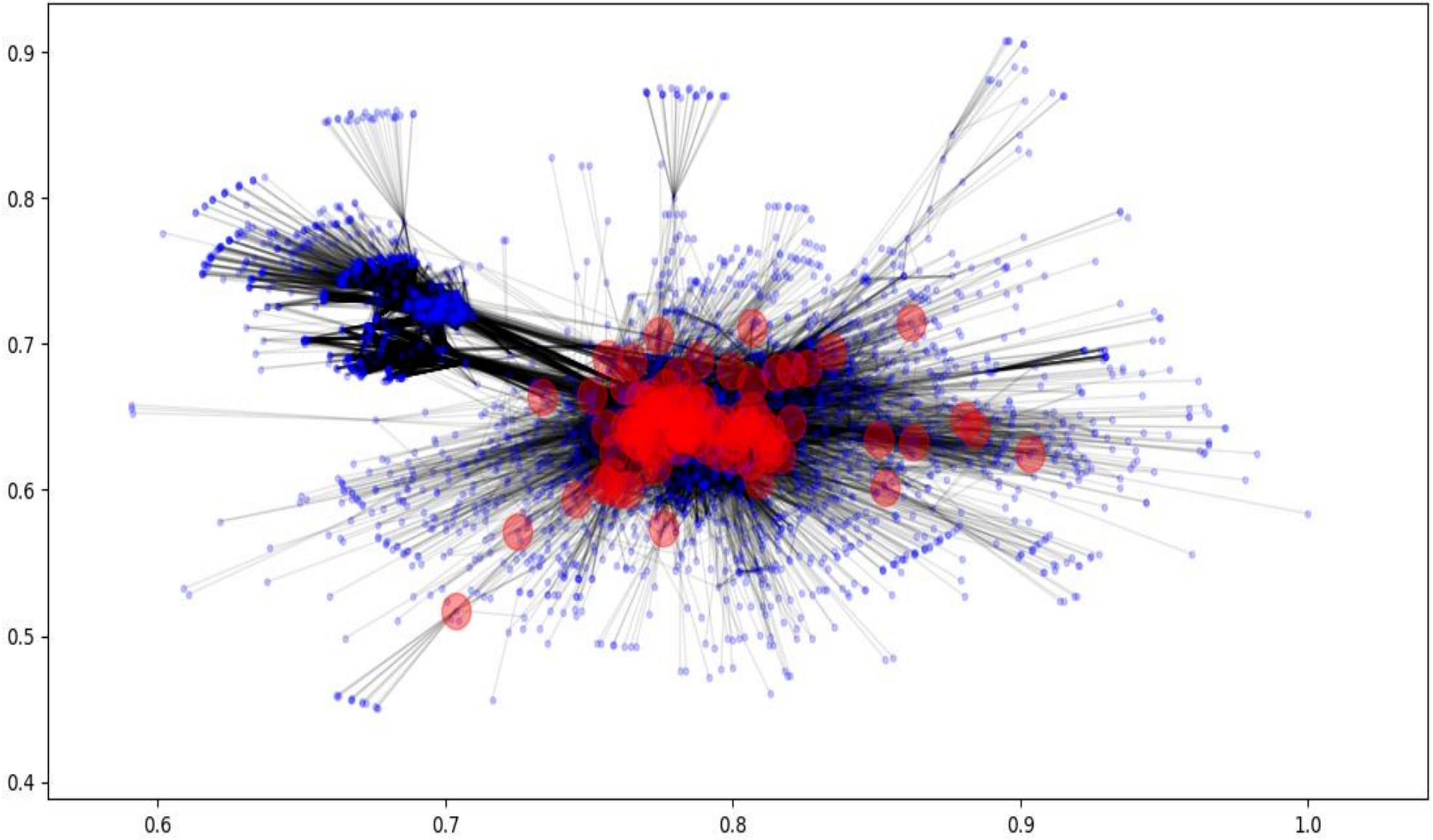
Among all **vulnerable products 50%** are Open Source [1)]




[1)] Based on the National Vulnerability Database (1999-2014) published by the National Institute of Standards and Technology

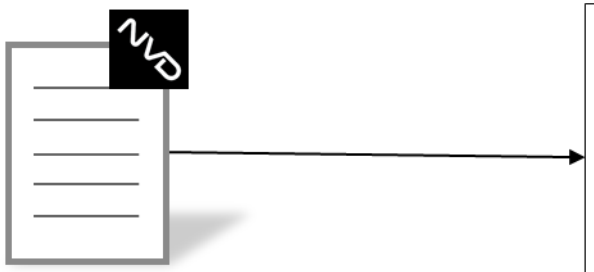
[2)] Source: Black duck software

© 2018 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC


0.9
0.8
0.7
0.6
0.5
0.4
0.6
0.7
0.8
0.9
1.0

***Repository***: https://github.com/some/project

***Commits***: 236AF0619, 1234E4D5C1A, ….



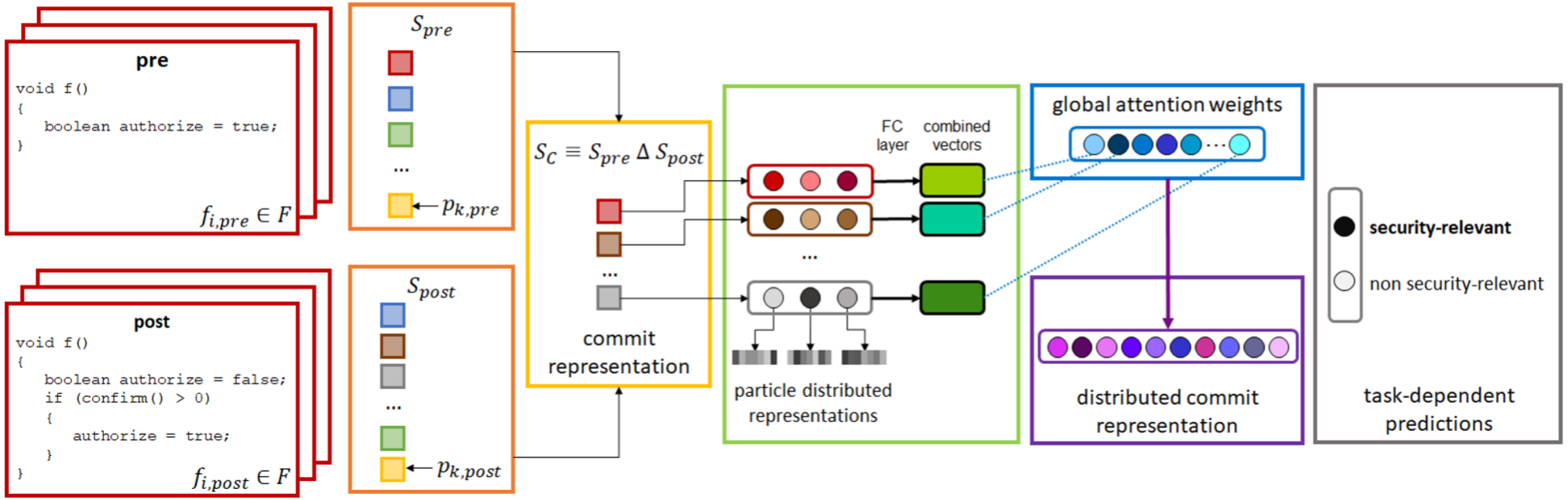

Vulnerability advisory (e.g., CVE)

Method x.y.z *ADDED*

Method a.b.c *MODIFIED*

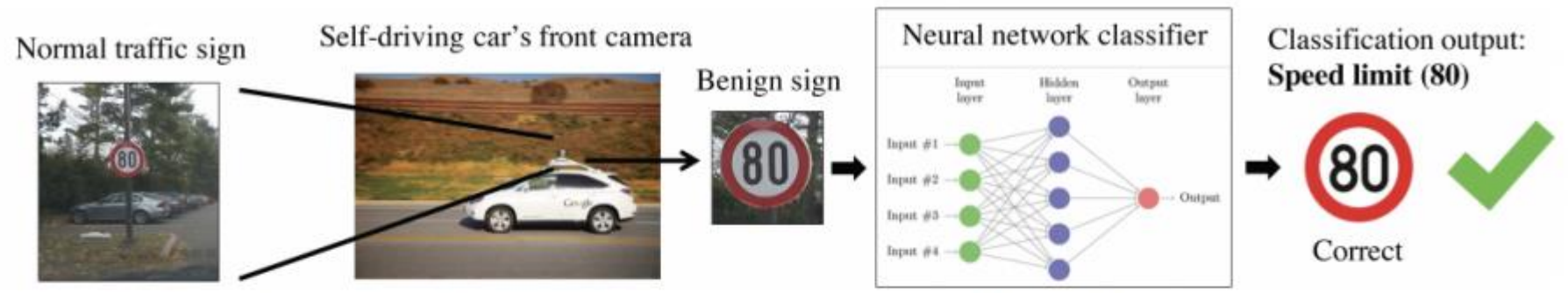
Method i.j.k *DELETED*

……

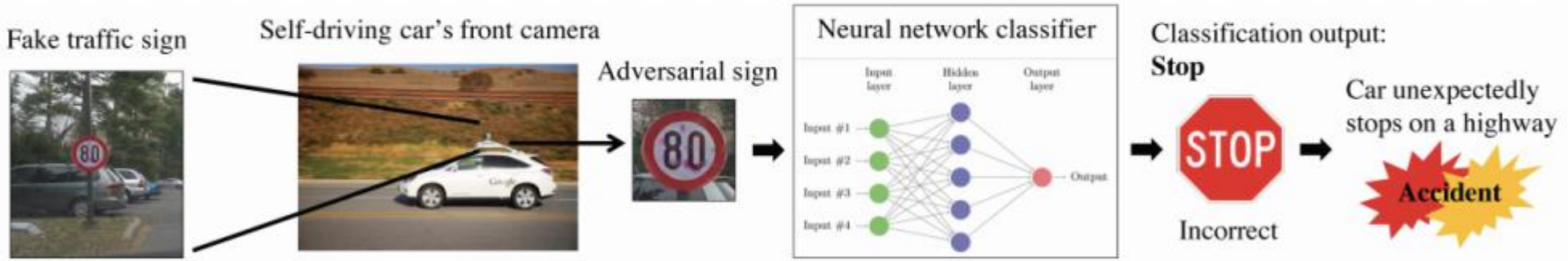



© 2018 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

# Attack to Machine Learning: Evasion Attack → Toxic Signs




(a) Operation of the computer vision subsystem of an AV under *benign conditions*



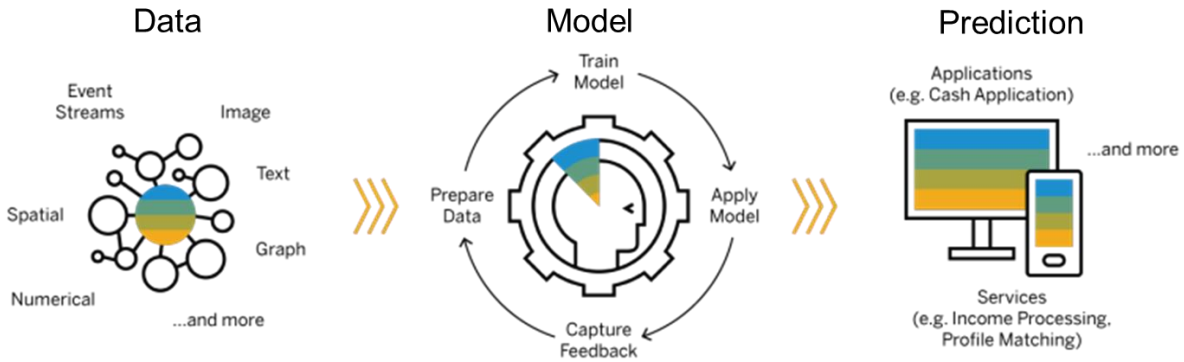

(b) Operation of the computer vision subsystem of an AV under *adversarial conditions*

Fig. 1. **Difference in operation of autonomous cars under benign and adversarial conditions.** Figure 1b shows the classification result for a drive-by test for a physically robust adversarial example generated using our Adversarial Traffic Sign attack.

*Sitawarin at al. 2018*

© 2018 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

# Security/privacy threats




| Property | Target | Attack | Mitigation |
|---|---|---|---|
| Confidentiality | Model | Duplicating model (IP) | Encryption (model & prediction) |
| Confidentiality | Model | Duplicating model (IP) | Watermarking |
| Privacy | Training data | Membership - Inference | DP-learning |
| Privacy | Training data | Membership - Inference | DP- Training data |
| Integrity | Prediction | Evasion | Testing data eng – |
| Integrity | Model | Poisoning | Training data control |
| Availability | System deployed | DoS | Redundancy - AC |

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I INTERNAL

# Conclusions

## Machine Learning for Security

- Intelligent algorithms for security problems (finding vulnerability, monitoring leaks)

## Security/Privacy for Machine Learning

- New attacks to machine learning (evasion, poisoning)
- Privacy for AI

© 2018 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

# Thank you.

Contact information:

**Michele Bezzi**
**Research Manager**
SAP Security Research

THE BEST RUN SAP®

SWAM

Logout

## List 624

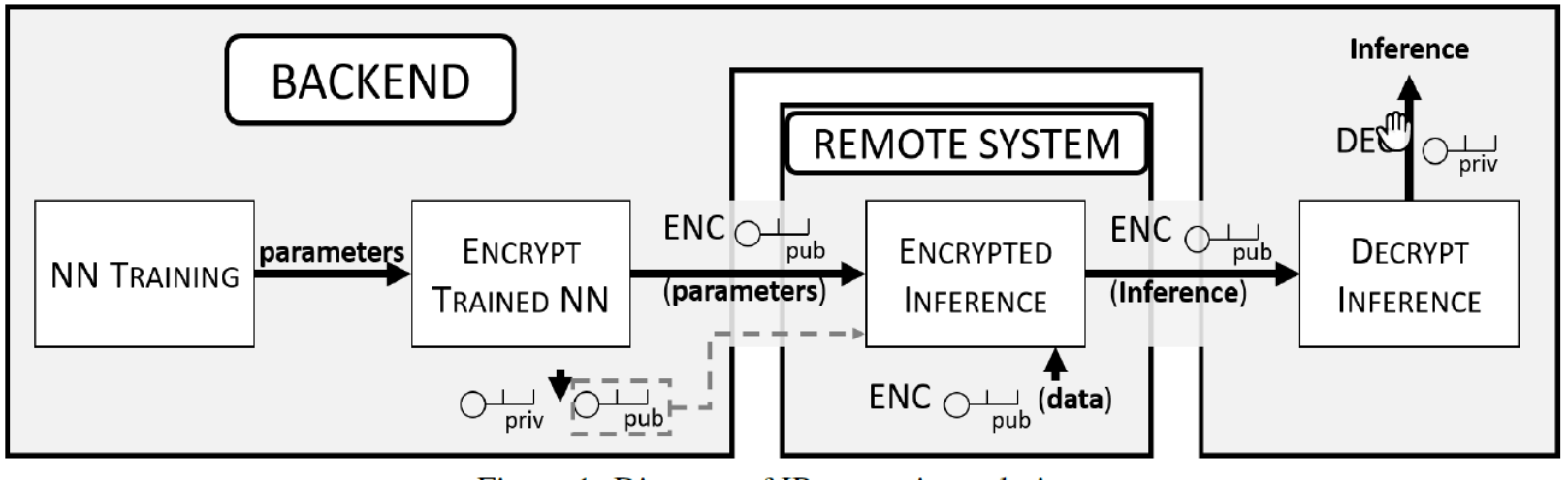
ALL | CODE | LEAK | LOG | PASSWORD | STREAMING | CHAT | WRITING | OTHER | GAMING

| Url | Date | Tags | Verified | Actions |
|---|---|---|---|---|
| https://pastebin.com/raw/TDx0cMSP | 11 Apr 2018 18:36 | code sap | ✖ | 👁 |
| https://pastebin.com/raw/heeefetH | 11 Apr 2018 17:35 | sap | ✖ | 👁 |
| https://pastebin.com/raw/ereRgU0k | 11 Apr 2018 15:41 | code sap | ✖ | 👁 |
| https://pastebin.com/raw/E1zwWLnn | 11 Apr 2018 12:28 | sap russian other | ✖ | 👁 |
| https://pastebin.com/raw/3U9GY0jd | 11 Apr 2018 10:16 | sap log | ✖ | 👁 |
| https://pastebin.com/raw/Ddr7FJ2s | 11 Apr 2018 10:12 | sap log | ✖ | 👁 |
| https://pastebin.com/raw/ViCtwVLR | 11 Apr 2018 9:9 | sap code | ✖ | 👁 |
| https://pastebin.com/raw/LGx1zqLw | 11 Apr 2018 5:38 | log sap log | ✖ | 👁 |
| https://pastebin.com/raw/fZMbpMkN | 11 Apr 2018 3:8 | code sap code | ✖ | 👁 |
| https://pastebin.com/raw/gtjTsrF1 | 10 Apr 2018 19:24 | sap leak log | ✖ | 👁 |
| https://pastebin.com/raw/M6fCErZu | 10 Apr 2018 18:37 | sap code | ✖ | 👁 |
| https://pastebin.com/raw/r7eTfYMU | 10 Apr 2018 18:25 | sap code | ✖ | 👁 |
| https://pastebin.com/raw/kwWMmmHr | 10 Apr 2018 18:8 | sap code | ✖ | 👁 |

SAP Web Assets Monitoring

© 2018 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

# Privacy for distributed neural networks




1. DNN Training (unencrypted data)
2. Encryption of trained DNN: Encrypt parameters → model can be distributed

3. inference, on decentralized systems, usi (protecting its IP)
4. Inference decryption: encrypted inference, to be decrypted only by the owner of the trained DNN

*Laurent Gomez, Alberto Ibarrondo, José Márquez, Patrick Duverger:* Intellectual Property Protection for Distributed Neural Networks - Towards Confidentiality of Data, Model, and Inference. ICETE (2) 2018: 313-320

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I INTERNAL