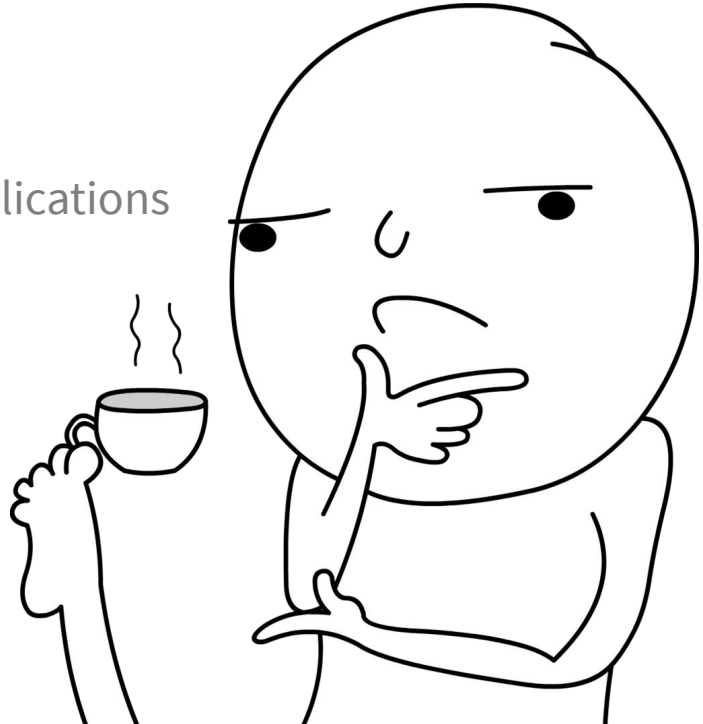
Trust, but Verify: Dangers of MaaS

Denis Kolegov, Antoniy Nikolaev
AISec Team

What it is All About?

What it is All About?

1. How to find different machine learning applications and frameworks on the Internet?
2. Are those solutions internet-wide secure?
3. What information can be obtained from them?
4. In which countries running machine learning applications can be found and what is their purpose?



Why?

Why?


MACHINE LEARNING
LIGHTYEAR
MACHINE LEARNING EVERYWHERE
imgflip.com

Cloud Text-to-Speech

Text-to-speech conversion powered by machine learning.

TRY IT FREE | VIEW DOCUMENTATION

face-api.js playground

Speech KITT

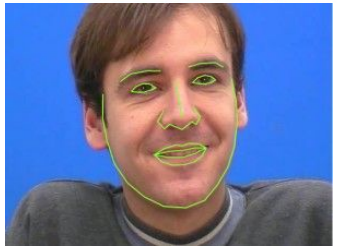
A flexible GUI for interacting with Speech Recognition

Speech KITT makes it easy to add a GUI to sites using Speech Recognition. Whether you are using annyang, a different library or webkitSpeechRecognition directly, KITT will take care of the GUI.

Speech KITT provides a gra for the user to start or stop Speech Recognition and see its also help guide the user on how to interact with your site using th uctions and sample commands. It can even be used to carry a na user, asking questions the user can answer with his voic questions.

OpenCV.js Demos

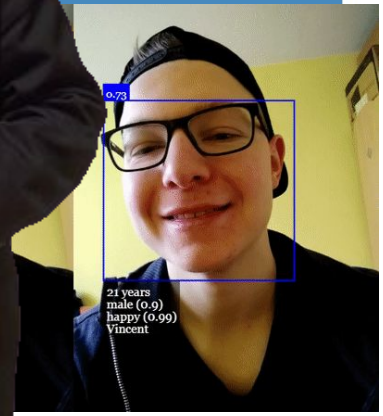
- Video processing (asm.js)
- Video processing (wasm)
- Face detection (asm.js)
- Face detection (wasm)



erested In

VOICE NG YOUR SITE

WITH OF CODE?



annyang! SpeechRecognition that just works

annyang is a tiny javascript library that lets your visitors control your site with voice commands.
annyang supports multiple languages, has no dependencies, weighs just 2kb and is free to use.

So What About Security*?

So What About Security*?

*(Not exactly)

The Problems

1. Many different interfaces of various ML frameworks are open and accessible from the Internet
2. Most of them don't have authentication mechanisms or got weak security policies
3. Default credentials almost the same and can be found in the documentation
4. Common vulnerabilities (web, software, etc.) both on server- and client-side

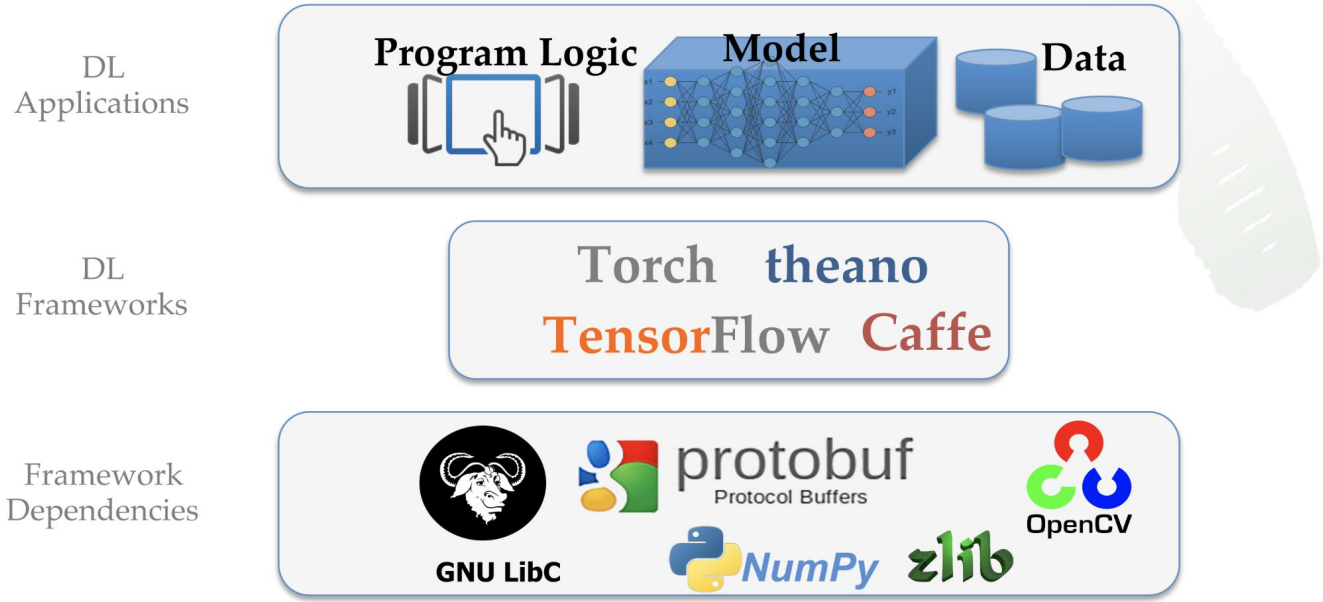
The Problems

1. Many diffe cessible from the Interne
2. Most of the curity policies
3. Default cre entation
4. Common v ent-side

Vulnerabilities

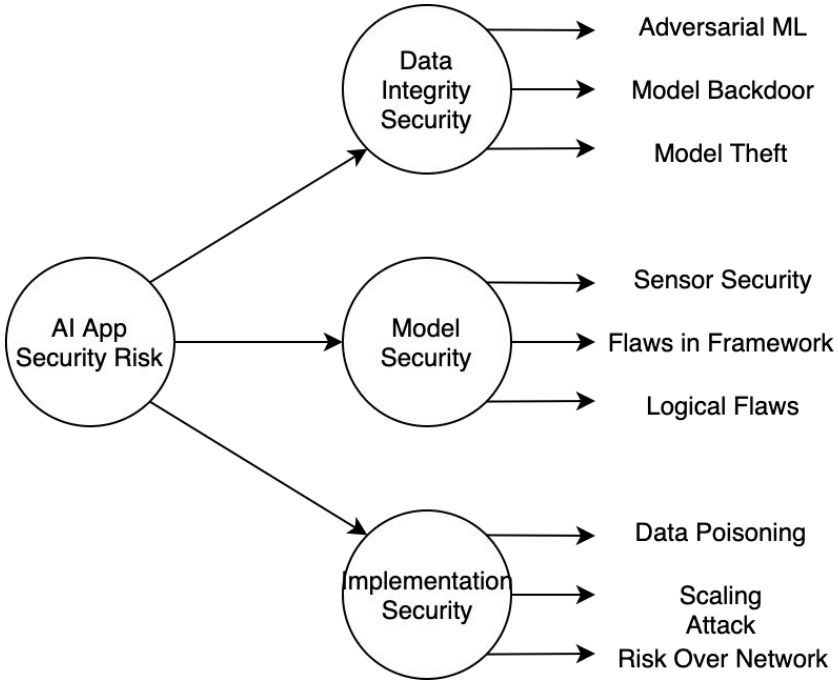
Machine Learning Frameworks

Software Layers in Machine Learning Apps



Li, K. (n.d.). *Exposing Vulnerabilities in Deep Learning Frameworks*.

AI or ML Application Threat Landscape



Li, K. (n.d.). *Reverse Engineering AI Models.*

Frameworks Complexity and Dependencies

ML Framework	Lines of Code	Number of Dep. Packages	Sample Packages
Caffe	127K+	137	Libprotobuf, libz, opencv, libopenblas
TensorFlow	887K+	97	numpy, librosa
Torch	590K+	48	xlua, qtsvg, opencv

Li, K. (n.d.). *Exposing Vulnerabilities in Deep Learning Frameworks*.

Common Bugs Found in ML Frameworks and Dependencies

ML Framework	dep. packages	CVE-ID	Potential Threats
TensorFlow	numpy	CVE-2017-12852	DOS
TensorFlow	wave.py	CVE-2017-14144	DOS
Caffe	libjasper	CVE-2017-9782	Heap Overflow
Caffe	openEXR	CVE-2017-12596	Crash
Caffe/Torch	opencv	CVE-2017-12597	Heap Overflow
Caffe/Torch	opencv	CVE-2017-12598	Crash
Caffe/Torch	opencv	CVE-2017-12599	Crash
Caffe/Torch	opencv	CVE-2017-12600	DOS
Caffe/Torch	opencv	CVE-2017-12601	Crash

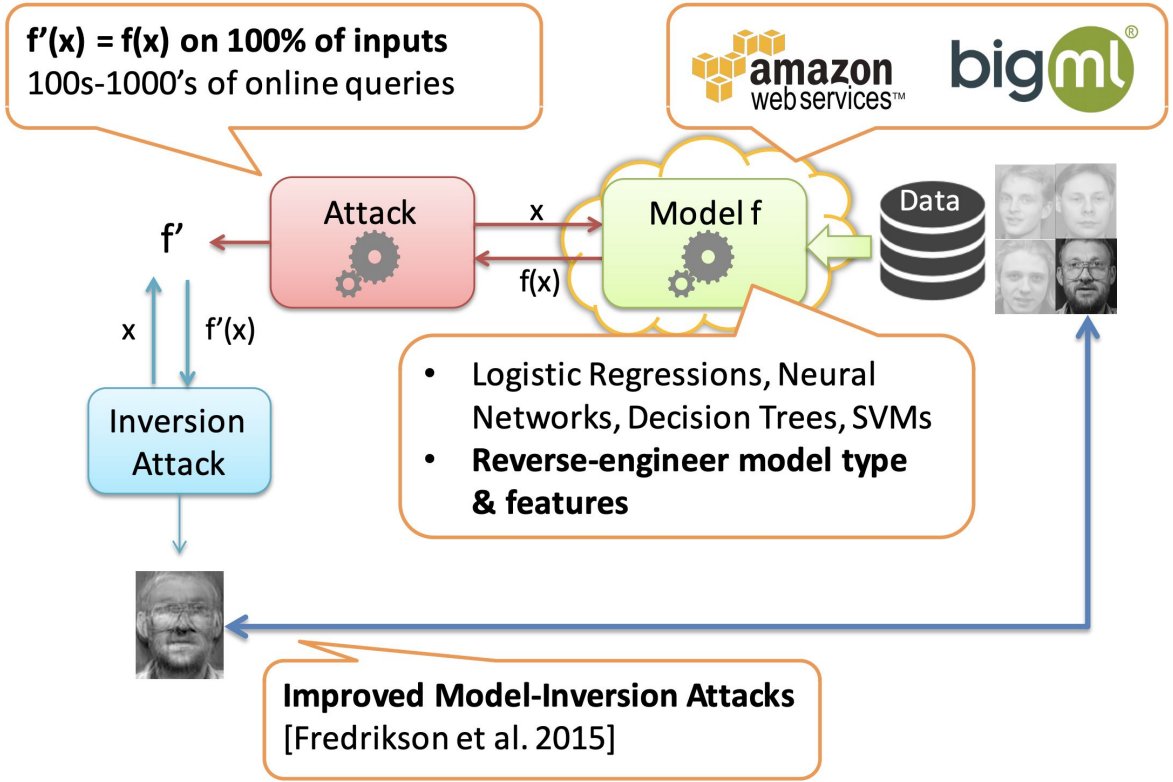
ML Framework	dep. packages	CVE-ID	Potential Threats
Caffe/Torch	opencv	CVE-2017-12602	DOS
Caffe/Torch	opencv	CVE-2017-12603	Crash
Caffe/Torch	opencv	CVE-2017-12604	Crash
Caffe/Torch	opencv	CVE-2017-12605	Crash
Caffe/Torch	opencv	CVE-2017-12606	Crash
Caffe/Torch	opencv	CVE-2017-14136	Integer Overflow

Li, K. (n.d.). *Exposing Vulnerabilities in Deep Learning Frameworks*.

What Does it Mean?

Machine Learning applications and related infrastructure (servers, wrappers, handlers) are vulnerable to different kinds of vulnerabilities: crashes, denial of service, integer and heap overflows, etc.

Model Extraction Attacks

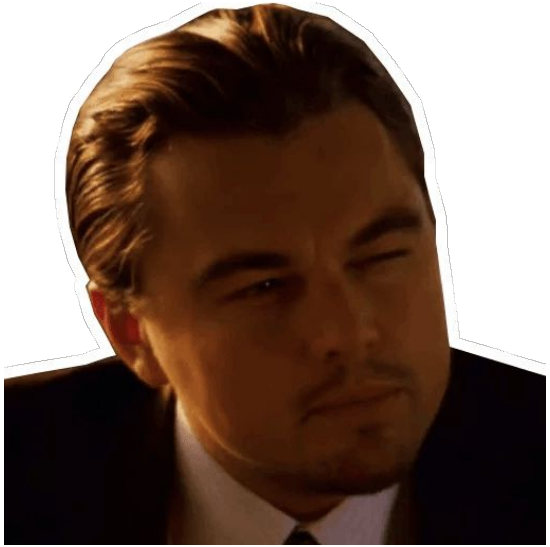


Tramèr, F. (2016). Stealing Machine Learning Models via Prediction APIs.

How to Find ML Frameworks and Applications?

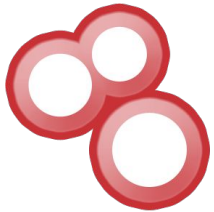
Search Engines

Search Engines



How can we search deeper?

Search Engines



SHODAN

censys

ZoomEye

Grinder Framework

grinder

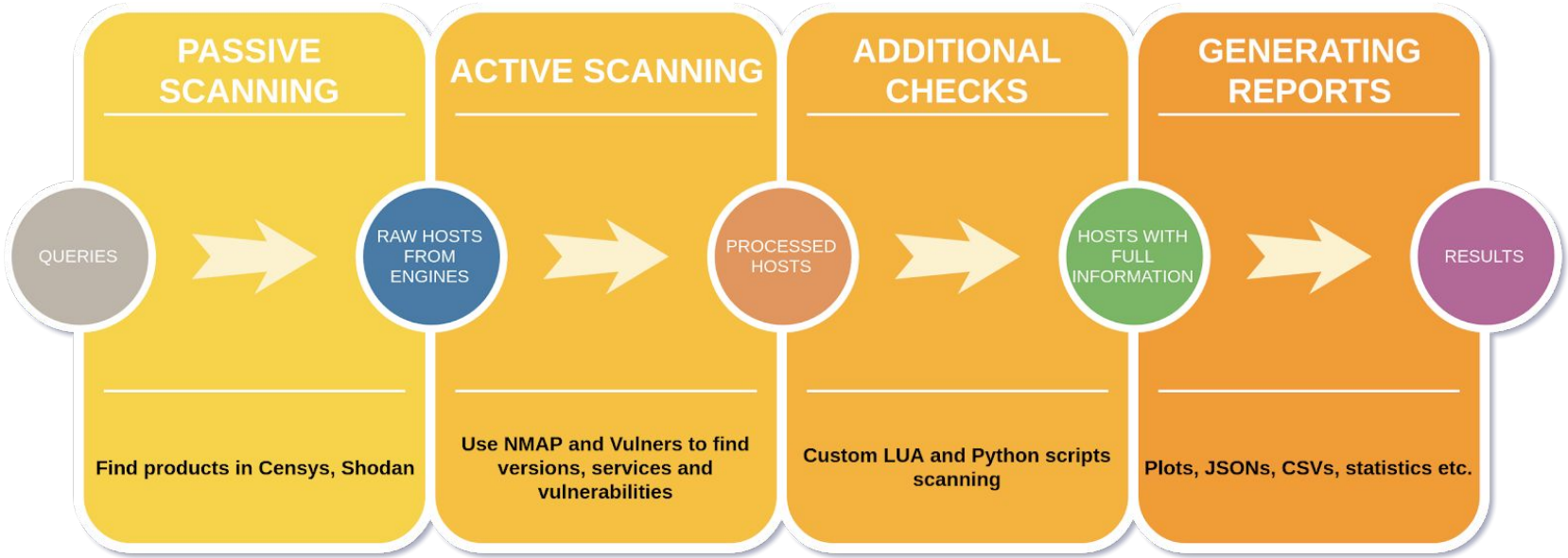
🔎 Python framework to automatically discover and enumerate hosts from different back-end systems (Shodan, Censys)

python · nmap · vulnerability-scanners · python-framework · shodan-api · vulners · censys-api

Python · GPL-2.0 · 4 · ★ 22 · 0 · 0 · Updated 7 days ago

github.com/sdnewhop/grinder

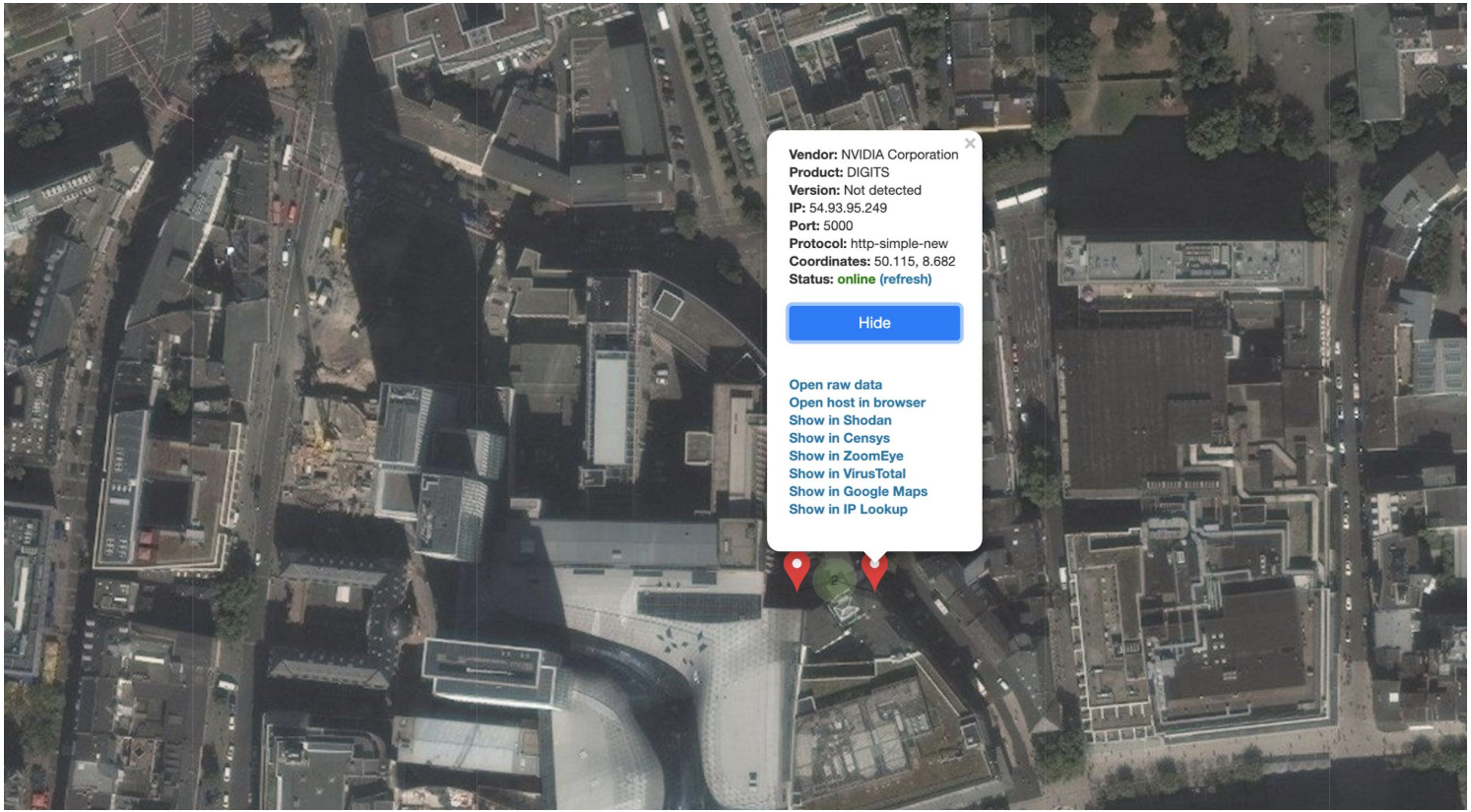
Grinder Workflow



Training Systems

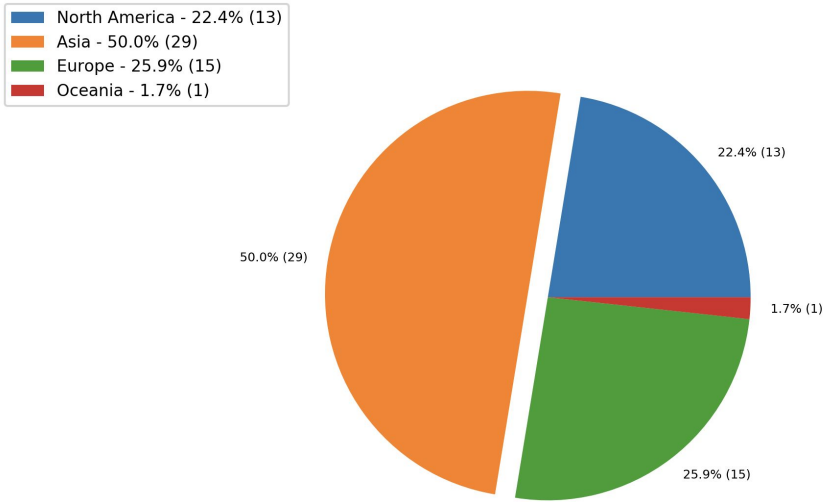
NVIDIA DIGITS

NVIDIA DIGITS

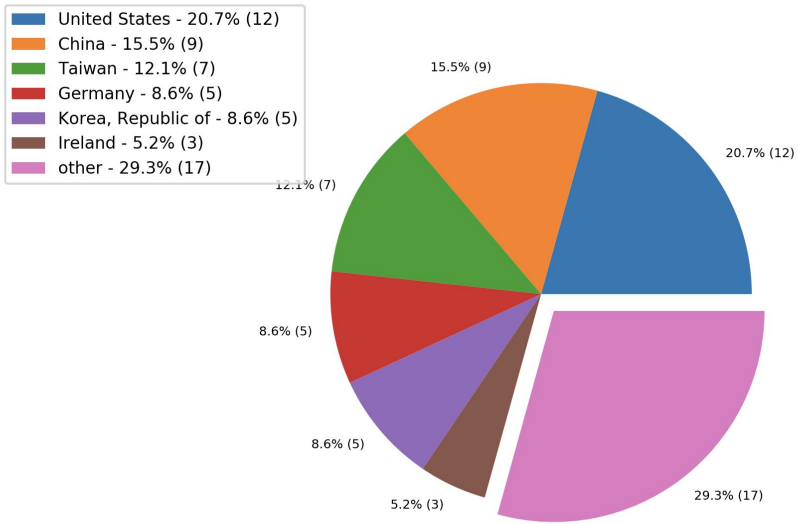


NVIDIA DIGITS

Percentage of nodes by continents

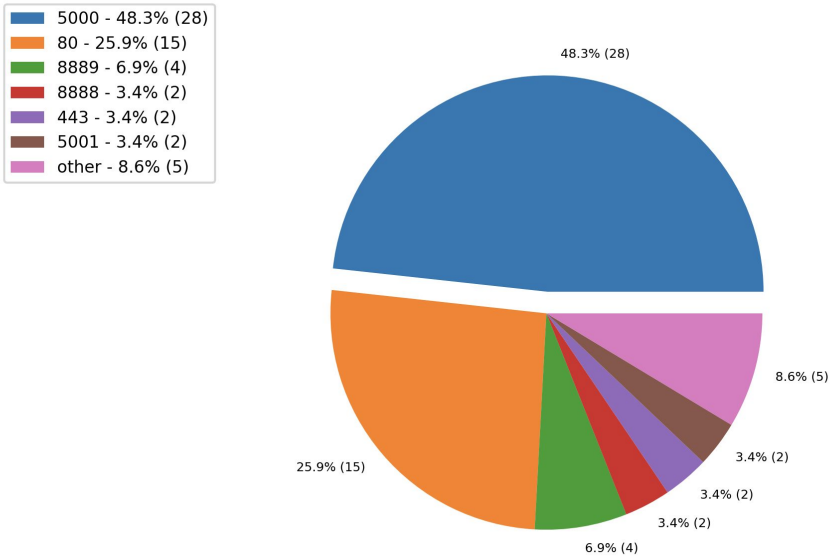


Percentage of nodes by countries

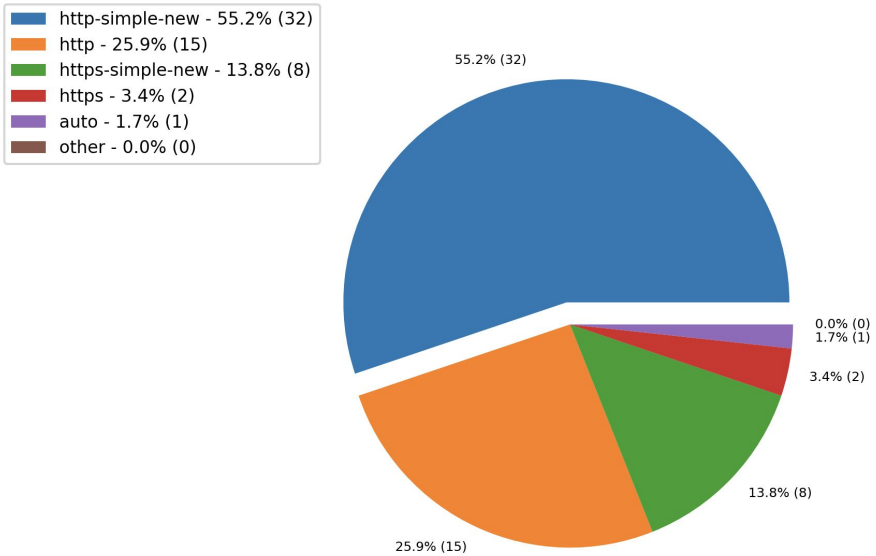


NVIDIA DIGITS

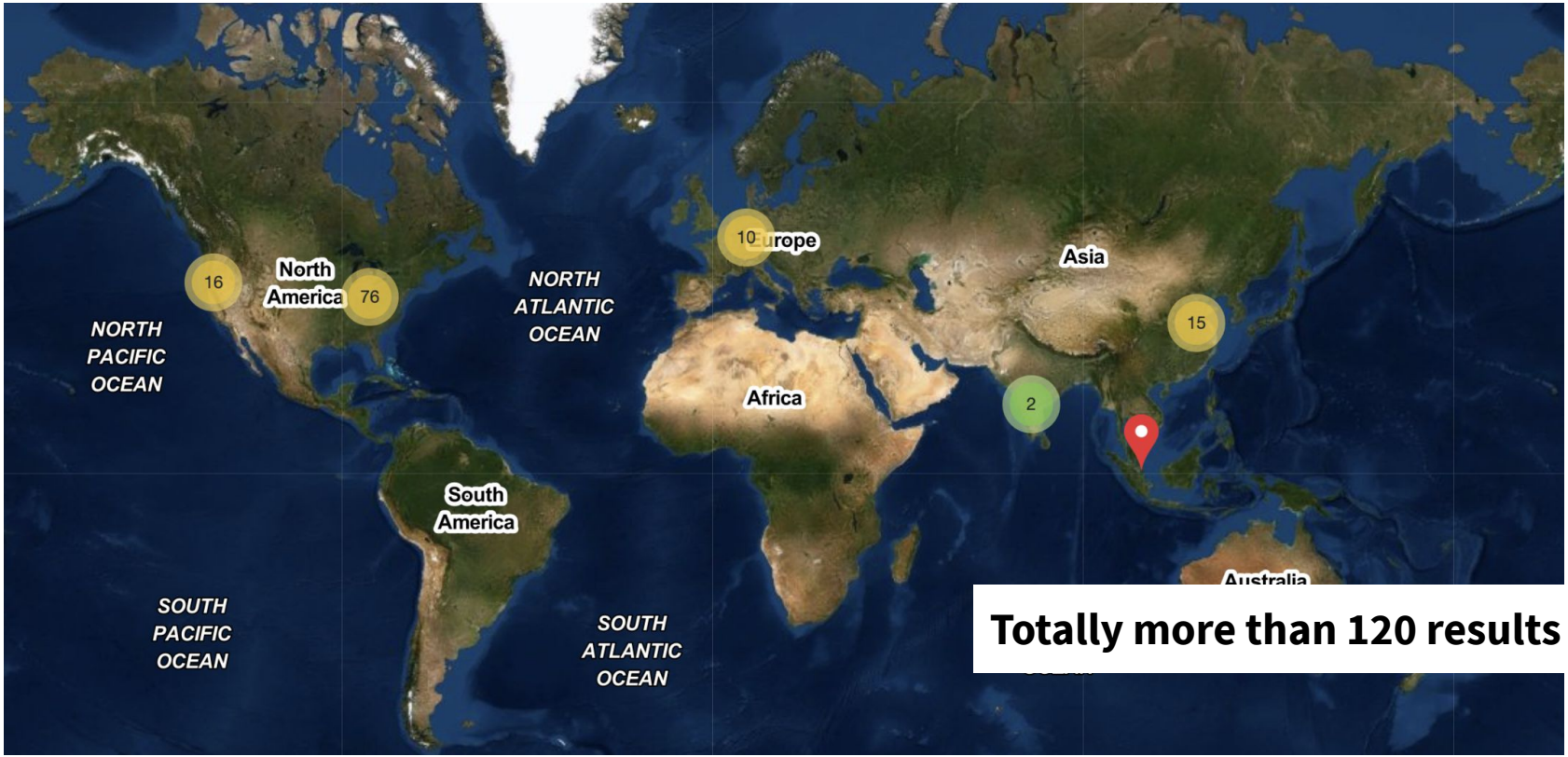
Percentage of nodes by ports



Percentage of nodes by protocols

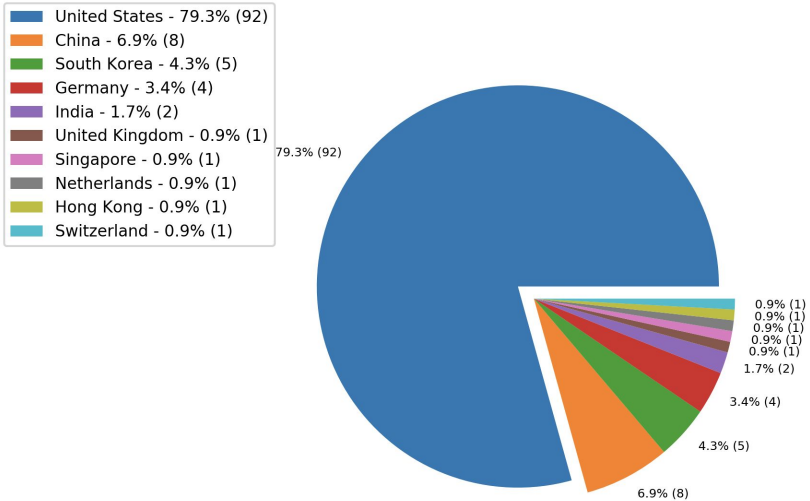


TensorBoard

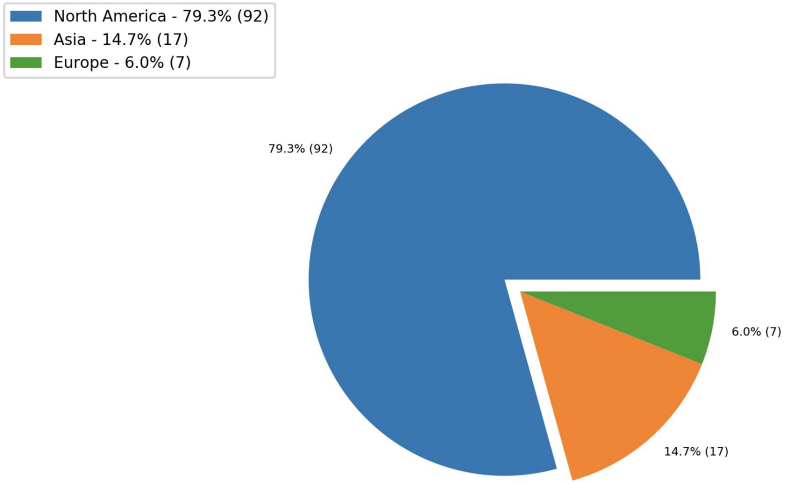


TensorBoard

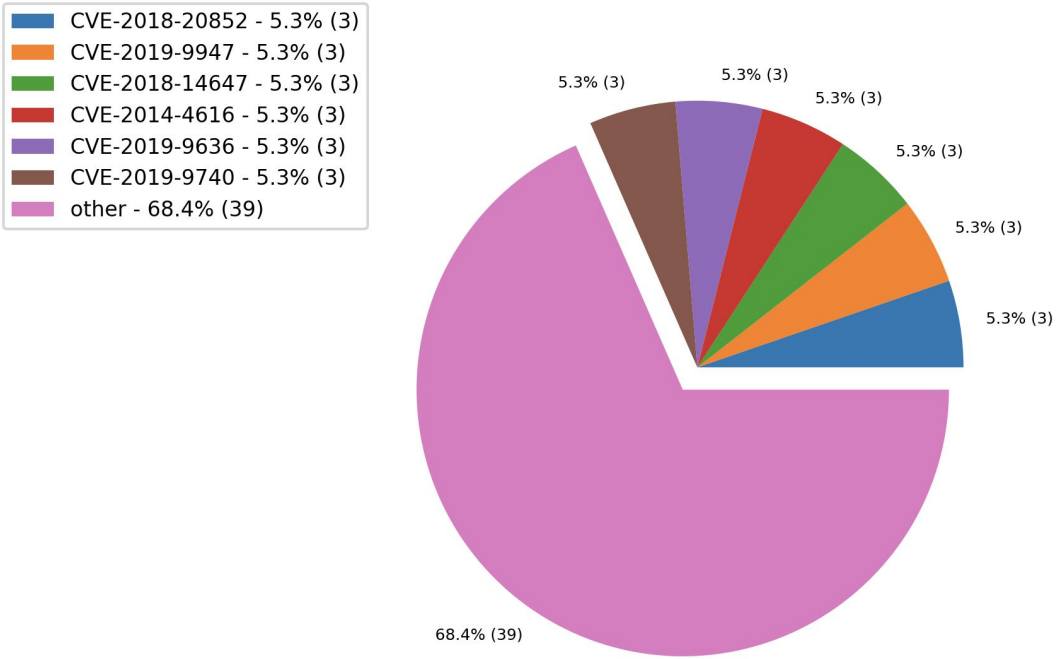
Percentage of nodes by countries



Percentage of nodes by continents

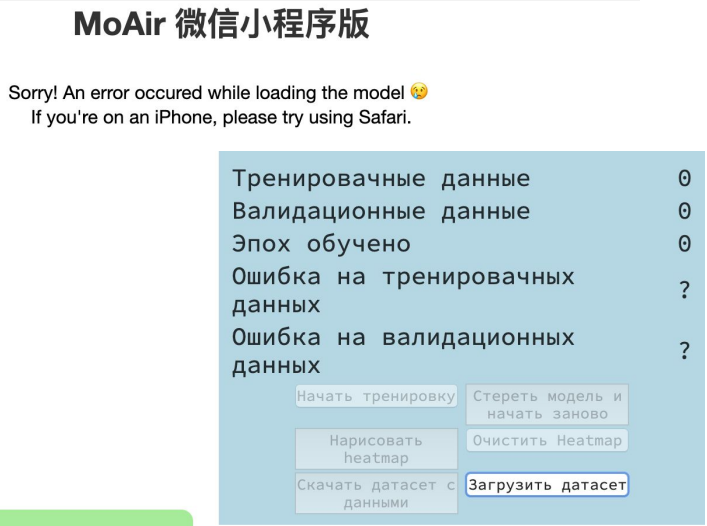


TensorBoard Vulnerabilities



JavaScript Applications

TensorFlow.js and Others



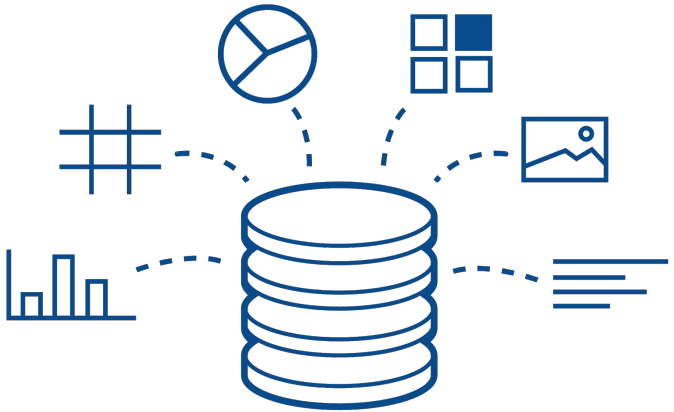
Your browser does not seem to support getUserMedia. 😭 This will probably only work in Chrome or Firefox.

Totally more than 1.000 different results were found

Open Databases with ML Data

Possible Impact

Sensitive Information Leakage: An attacker can gain full access to sensitive information about the system, infrastructure, processes, and system paths. Moreover, the logs may contain personal and private information.



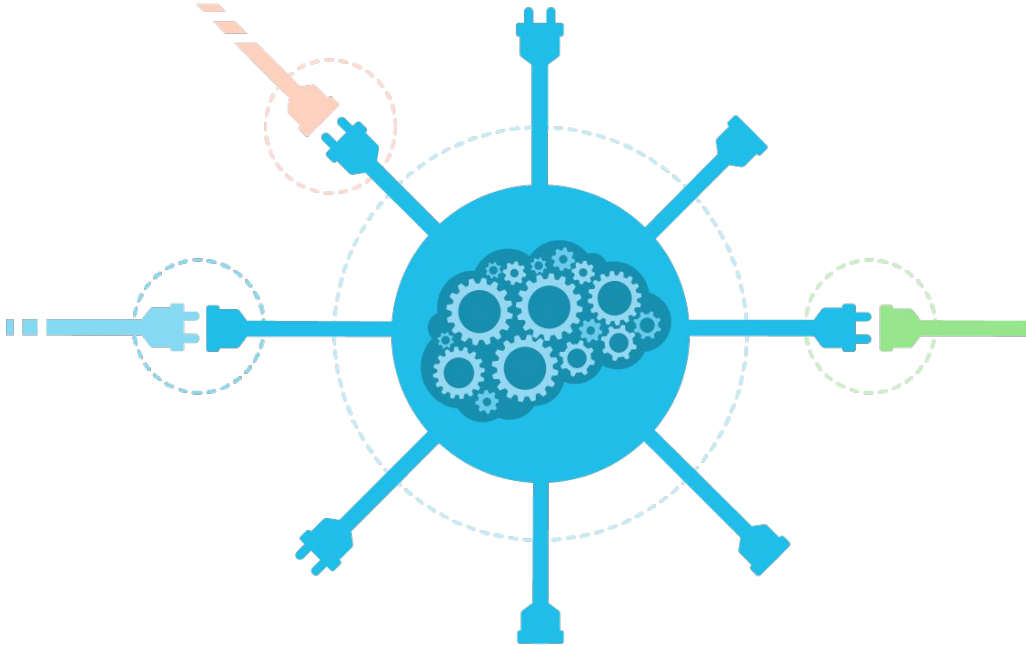
Possible Impact

Information Theft: an attacker can gain access to datasets, tags, and labels with different databases and collections

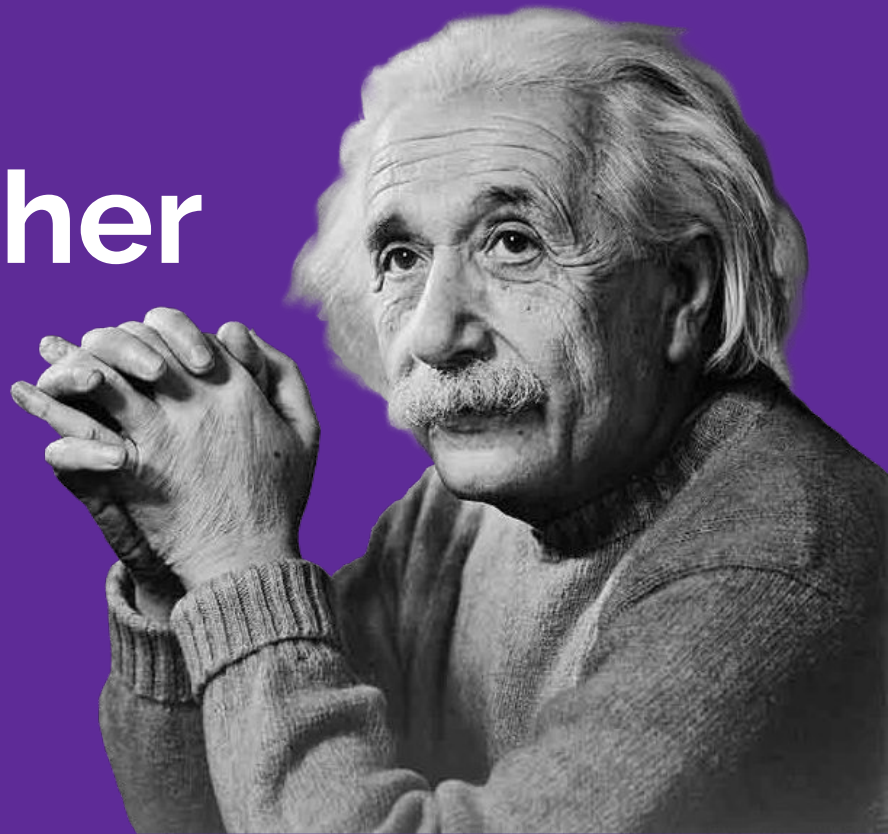
Running Containers

Possible Impact

API Mechanisms Exposure: an attacker can manage remote containers and images; run, delete or stop them

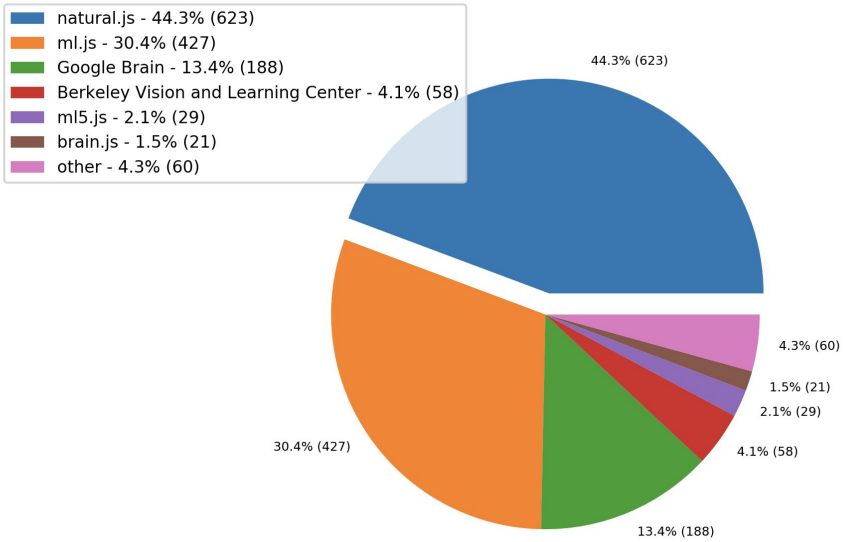


What About Other Applications?

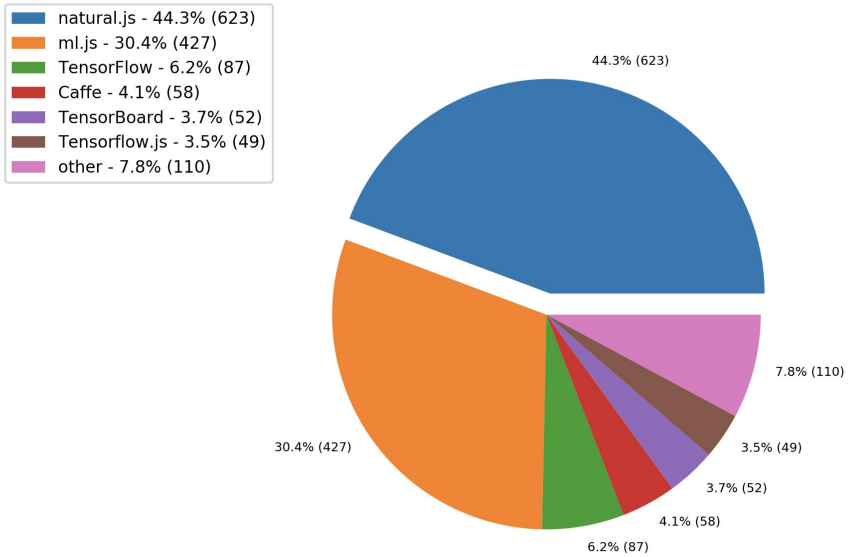


Results (July 2019)

Percentage of nodes by vendors

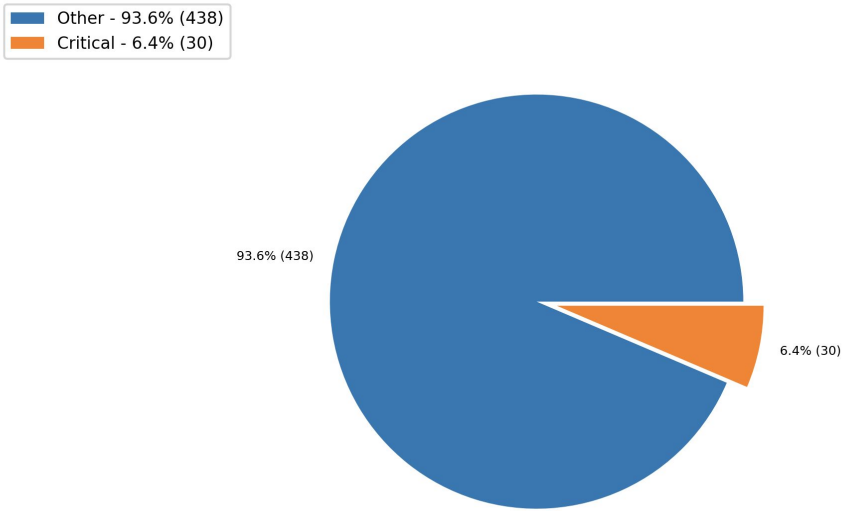


Percentage of nodes by products

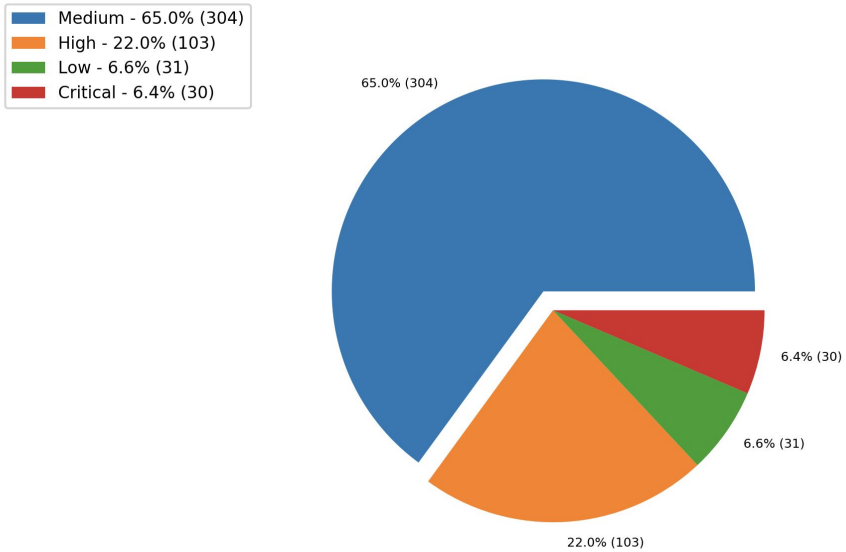


Vulnerabilities (July 2019)

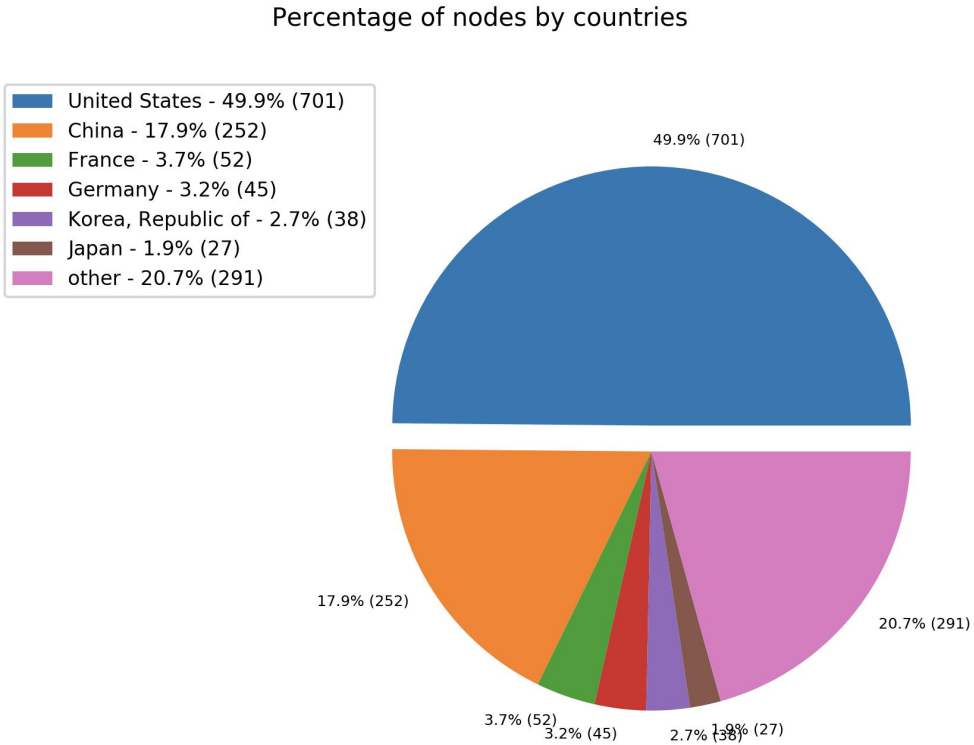
Percentage of critical vulnerabilities



Percentage of vulnerabilities by CVSS v3.0 rating



Distribution by Countries



Results Map


4
5
241
Europe
Asia
North
NORTH
ATLANTIC
OCEAN
NORTH
PACIFIC
OCEAN
353
12
2
Africa
25
South
America
9
Australia
26
SOUTH
PACIFIC
OCEAN
SOUTH
ATLANTIC
OCEAN
INDIAN
OCEAN

AIFinger Project

The goals of the project is to provide tools and results of passive and active fingerprinting of Machine Learning Frameworks and Applications using a common Threat Intelligence approach and to answer the following questions:

- How to detect ML backend systems on the Internet and Enterprise network?
- Are ML apps secure at Internet scale?
- What is ML apps security level in a general sense at the present time?
- How long does it take to patch vulnerabilities, apply security updates to the ML backend systems deployed on the Internet?

sdnewhop.github.io/AISec/

github.com/sdnewhop/AISec

Contributors:

- Sergey Gordeychik
- Anton Nikolaev
- Denis Kolegov
- Maria Nedyak

AIFinger Project Coverage

- Frameworks
  - TensorFlow
  - NVIDIA DIGITS
  - Caffe
  - TensorBoard
  - Tensorflow.js
  - brain.js
  - Predict.js
  - ml5.js
  - Keras.js
  - Figue.js
  - Natural.js
  - neataptic.js
  - ml.js
  - Clusterfck.js
  - Neuro.js
  - Deeplearn.js
  - Convnet.js
  - Synaptic.js
  - Apache mxnet
- Databases with ML Content
  - Elasticsearch with ML data
  - MongoDB with ML data
  - Docker API with ML data
- Databases
  - Elasticsearch
  - Kibana (Elasticsearch Visualization Plugin)
  - Gitlab
  - Samba
  - Rsync
  - Riak
  - Redis
  - Redmon (Redis Web UI)
  - Cassandra
  - Memcached
  - MongoDB
  - PostgreSQL
  - MySQL
  - Docker API
  - CouchDB
- Job and Message Queues
  - Alibaba Group Holding AI Inference
  - Apache Kafka Consumer Offset Monitor
  - Apache Kafka Manager
  - Apache Kafka Message Broker
  - RabbitMQ Message Broker
  - Celery Distributed Task Queue
  - Gearman Job Queue Monitor
- Interactive Voice Response (IVR)
  - ResponsiveVoice.JS
  - Inference Solutions
- Speech Recognition
  - Speech.js
  - dictate.js
  - p5.speech.js
  - artyom.js
  - SpeechKITT
  - annyang

… and many more

Thanks for Attention.
Got Questions?

@dnkolegov

@manmoleculo

github.com/sdnewhop/grinder

sdnewhop.github.io/AISec/