HSLU Lucerne University of Applied Sciences and Arts

SecTalk

Studiengang Information & Cyber Security

Offensive Security im Fokus

OT und IT

Major

Attack Specialist & Penetration Tester

Information Technology

22 November 2023

FH Zentralschweiz

Begrüssung der Studiengangsleitung

Bernhard M. Hämmerli

Willkommen in der HSLU-I SecTalk Serie

- Wir freuen über Ihr Interesse und Ihren Besuch bei uns.
- Sie sind am Ort des Geschehens für angewandte Cybersecurity.
- BSc ICS seit 2018: Heute grösster Lehrgang der Hochschule ca. 320 Studierende. 6000 Studierendenstunden nur für ICS!
  ➔ Studierenden sind cool
  ➔ Studium ist cool
  ➔ Treffe Top Security Experten
- Baue deine Security-Zukunft und komme zu uns: es lohnt sich.
- Neu Master Information & Cyber Security MSc ICS an 2025
  Sie können sich in der Interessentenliste eintragen:
  Umfrage Interesse am Masterstudium ICS (Fortsetzung des BSc ICS) (office.com)

- Frauenevents: **7. Dezember**, zwischen 18:00 und 20:30 Uhr **Bitfee AG (Hammergut 1, 6330 Cham)**
  Frauen, bitte meldet euch bis zum 4. Dezember über den Anmeldelink an.

Information & Cyber Security Generalist

- Die HSLU–Informatik bildet Security Generalisten in fünf Richtungen aus:

- Information Security Technology
- Information Security Management
- Digital Forensics & Incident Response
- Attack Specialist & Penetration Tester
- Security & Cloud Technologies
- Free Choice

ca. 100 unterschiedliche Job Profile bieten für jeden Geschmack, jede Richtung und jedes Temperament ein passendes Job-Profil

Peter Infanger
peter.infanger@hslu.ch
Verantwortlicher Major MST

Sebastian Obermeier
sebastian.obermeier@hslu.ch
Verantwortlicher Major MSP
Director Critical Infrastructure Security Lab

Hannes Spichiger
hannes.spichiger@hslu.ch
Verantwortlicher Major MSF

Major MSM – Information Security Management (8 von 11 auswählen)

- ASTAT: Applied Statistics for Data Science
- CISO ISSUES: CISO Issues
- CRS: Crisis Recovery Strategies
- DLP: Data Leakage Protection
- DSO: Datenschutz in Organisationen
- CYBERCRIME: aka Forensic Readiness
- HOA: Human and Organizational Aspects of Information Security
- ADRM: Advanced Risk Management
- KRKO: Krisenmanagement & -kommunikation
- SGC: Secure Governance and Compliance
- SGPOR: Secure Business Processes in Organizations

Major MST – Information Security Technology (8 von 12 auswählen)

- ASTAT: Applied Statistics for Data Science
- CYBER: Cyber Defense
- DLP: Data Leakage Protection
- KRINF: Kritische Infrastruktur Sicherheit
- KRINFLAB: Kritische Infrastruktur Labor
- MOBILSEC: Mobile Security
- NETDA: Network Defense & Architecture
- REVE1: Reverse Engineering 1
- REVE2: Reverse Engineering 2
- SIOT: Secure IoT
- SOC: Security Operation Center Issues
- SYSSEC: System & Security

Major MSF – Digital Forensics & Incident Response (8 von 12 auswählen)

- CF: Computer Forensics
- CRS: Crisis Recovery Strategies
- CYBER: Cyber Defense
- DFF: Digital Forensic Foundation
- CYBERCRIME: aka Forensic Readiness
- HTCLAW: High Tech Cybercrime & Law
- IRFORENSIC: Angewandte Incident Response und IT Forensik
- MALWLAB: Malware Analysis Lab
- SOC: Security Operation Center Issues
- SYSSEC: System & Security
- MOBINFSEC: 5G Mobile Networks, Technologies & Security
- MCIP: Mobile, Cloud & IoT Forensics

Major MSP – Attack Specialist & Penetration Tester (8 von 12 auswählen)

- ADPENTEST: Advanced Penetration Testing
- BDM: Big Data Management
- CYBER: Cyber Defense
- CYBER2: Cyber 2 - Advanced Penetration Testing and Bug Bounties
- KRINF: Kritische Infrastruktur Sicherheit
- KRINFLAB: Kritische Infrastruktur Labor
- MALWLAB: Malware Analysis Lab
- ML: Machine Learning
- REVE1: Reverse Engineering 1
- REVE2: Reverse Engineering 2
- SOC: Security Operation Center Issues
- SYSSEC: System & Security

Major MSC – Security & Cloud Technologies (8 von 12 auswählen)

- CAB: Cloud-Services Angebot & Betrieb
- CI: Cloud Infrastructure
- CSARCH: Cyber Security & Architecture
- CT: Cloud Technology
- CYBER: Cyber Defense
- CYBERCRIME: aka Forensic Readiness
- IRFORENSIC: Angewandte Incident Response und IT Forensik
- ITIA: IT Infrastructure Automation
- NETDA: Network Defense & Architecture
- SOC: Security Operation Center Issues
- SYSSEC: System & Security
- MCIP: Mobile, Cloud & IoT Forensics

Major Attack Specialist & Penetration Testing
8 aus 12 Modulen zu wählen

- ADPENTEST - Advanced Penetration Testing
- BDM - Big Data Management
- CYBER - Cyber Defense
- CYBER2 - Cyber 2 Advanced Penetration Testing & Bug Bounties
- KRINF - Kritische Infrastruktur Sicherheit
- KRINFLAB - Kritische Infrastruktur Labor
- MALWLAB - Malware Lab
- ML - Machine Learning
- REVE1 - Reverse Engineering 1
- REVE2 - Reverse Engineering 2
- SOC - Security Operation Center Issues
- SYSSEC - System & Security

Sebastian Obermeier
sebastian.obermeier@hslu.ch
Verantwortlicher Major MSP

Um Sicherheitslücken in existierenden Systemen zu identifizieren ist es essenziell, das Denken und Vorgehen von Angreifern zu verstehen.
Neben Enterprise-Systemen behandelt der MSP-Major kritische Infrastrukturen und nutzt das HSLU eigene Labor um im praktischen Umfeld zu vermitteln, wie Unterstationen im Energiebereich angegriffen und verteidigt werden können.

Offensive Security
Motivation

- Schwachstellenidentifikation
- Verständnis des Angreifers
- Compliance und Sicherheitsbewertungen

Offensive Security ist unerlässlich in einer ganzheitlichen Cyber Strategie

Wenn die Fabrik denkt sie sei im Silicon Valley:

OT-Sicherheit ist mehr als nur ein IT-Upgrade im Overall

HSLU Lucerne University of Applied Sciences and Arts

Agenda

Begrüssung, Eröffnung
Prof. Dr. Bernhard Hämmerli, Studiengangsleiter Information & Cyber Security, HSLU
Prof. Dr. Sebastian Obermeier, Information & Cyber Security HSLU, Director Critical Infrastructure Lab, Hochschule Luzern

Cybersecurity in the Era of AI
Nico Heise, Solution Architect, Cybersecurity Services IBM Consulting

Pentesters Diary – Geschichten aus dem Alltag eines Penetration Testers
Yves Kraft, Head of Cyber Security Academy, Oneconsult AG

IT ist anders … OT auch!
Markus Lenzin, Mitinhaber und Senior Cyber Security Consultant für kritische Infrastrukturen, ALSEC Cyber Security Consulting AG

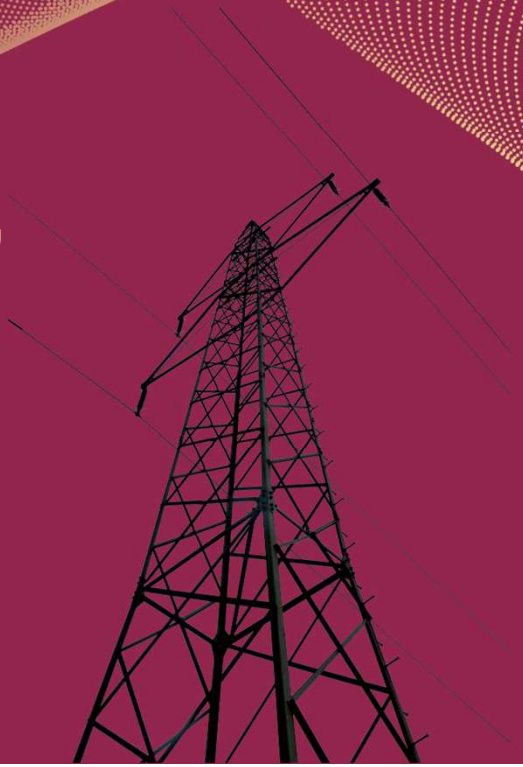
Apéro riche

HSLU Hochschule Luzern

Information &
Cyber Security Tech-Meeting

ALSEC CYBERSECURITY CONSULTING AG

Cyber Security,
"IT ist anders – OT auch!"



ALSEC Cyber Security Consulting AG

Rotkreuz, 22. September 2023

AGENDA „IT ist anders - OT auch!“

- Referent
- Bedrohungen
- Politische und gesetzliche Lage
- Input NCSC zur Meldepflicht
- IT und… oder… ohne… mit… OT….
- KRINFLAB
- Fragen & Antworten

Referent

Markus LENZIN

- Dozent an der HSLU / VSE für Cyber Security in der OT
- +30 Jahre Erfahrung im Service Management und Projekten wie Systemen, Netwerken und Telecommunications im SCADA Umfeld.
- +30 Jahre Erfahrung im Betrieb von kritischer IT und OT Infrastruktur im Energiesektor.
- Mitglied in nationalen und internationalen Arbeitsgruppen bezüglich Systemen, Applikationen und Cyber Security in IT und OT.

Branche: General IT, OT in Energie und Industrie (kritische Infrastrukturen)

ALSEC Portfolio

Consulting
- Security Strategy
- Security Consulting
- Security Product Evaluation
- Security Supplier Management

Audits
- Security Assessments
- Security Audits

Services
- CISO as a Service
- Rent an OT Labor
- Security Operation / Monitoring

Support
- Betriebsunterstützung
- Projektunterstützung
- Incident Management

Schulung
- Awareness/Training/Education
- Engineers / Architects
- Führungskräfte
- Verwaltungsräte

OT Labor
- Training im OT Labor
- Evaluation / POC von OT Geräten

ALSEC ACADEMY

Schulung / Training

VSE Certified OT Security Engineer

OT Labor HSLU

Awareness OT- / IT-Security

Management / Mitarbeiter

OT- / IT-Security Kurse

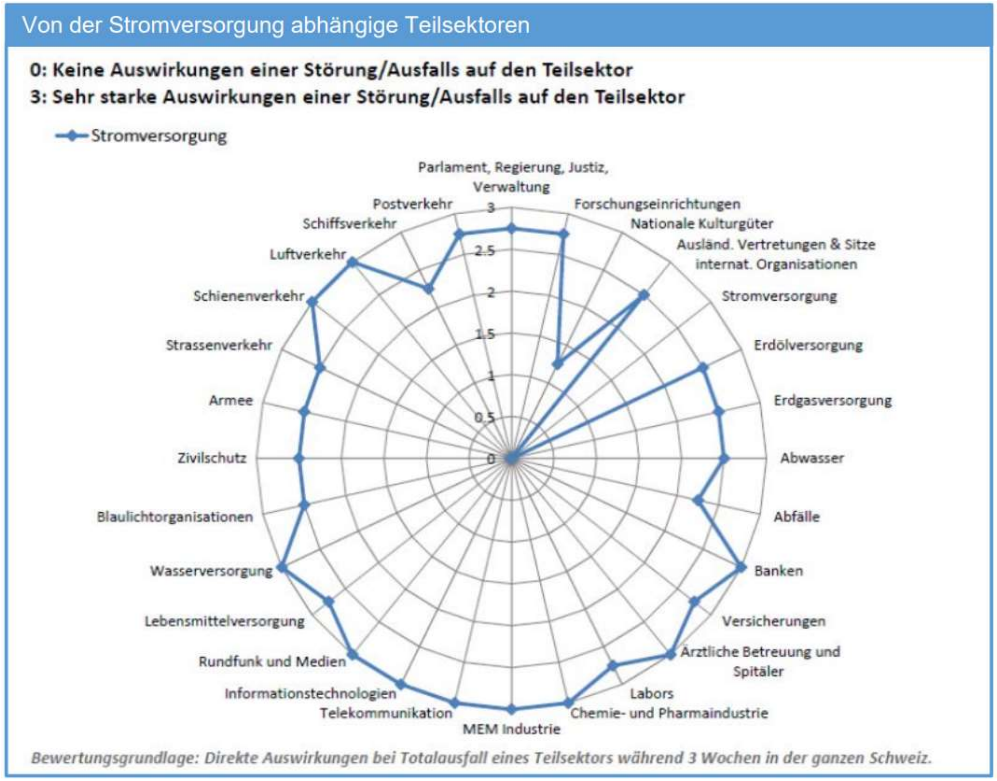
Spezialisten / Führungskräfte / Sicherheitsverantwortliche

Studies OT-Security

Spezialisten / Studenten

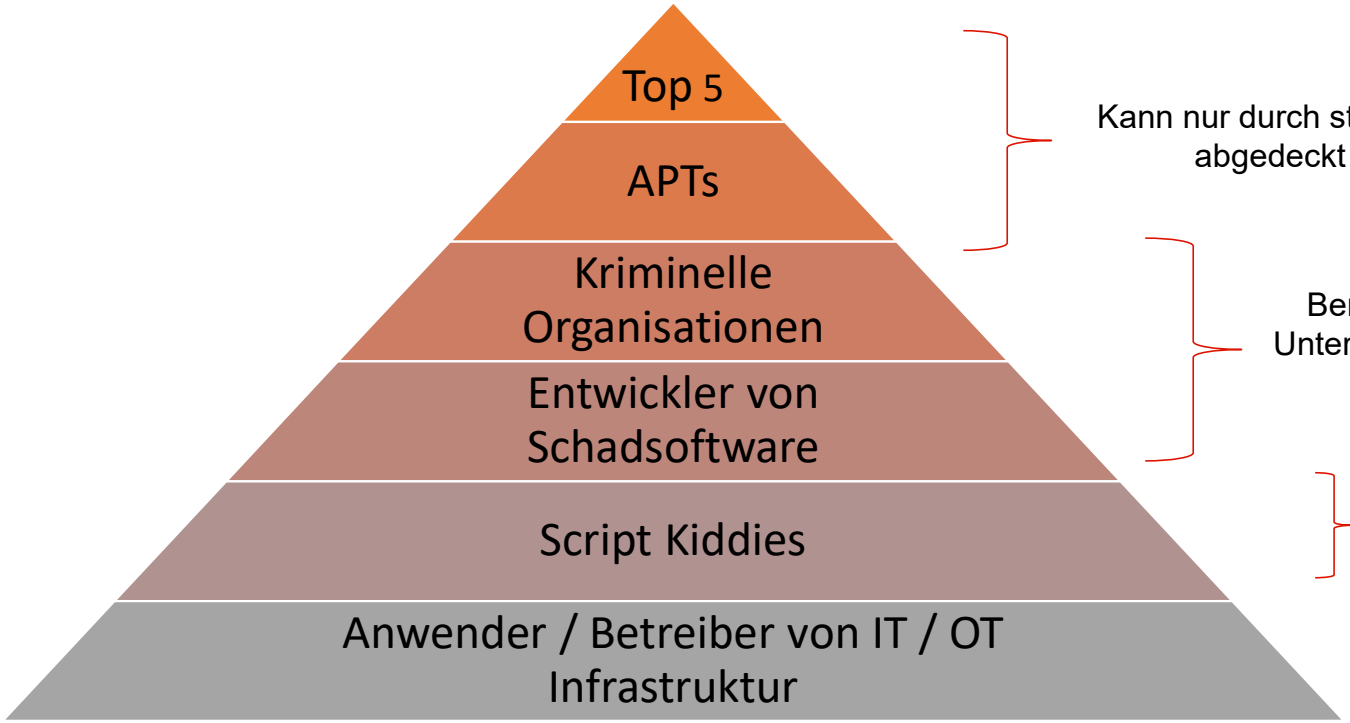
Bedrohungen

Kritikalität des Energiesektors



Source: Federal Office of Civil Protection

Akteure im Cyberraum

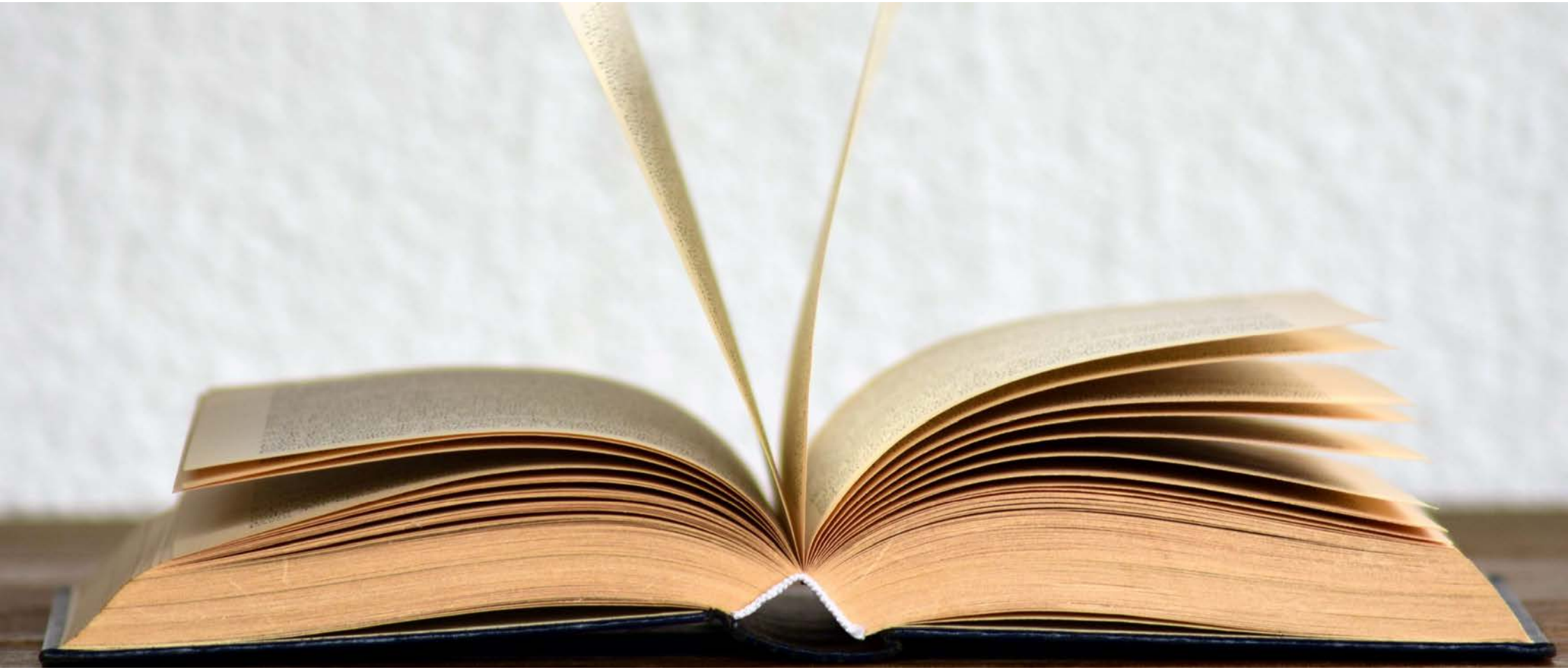


Bedrohungen für «Industrial Control Systems»

Top 10 Bedrohungen	Trend seit 2019
Einschleusen von Schadsoftware über Wechseldatenträger und mobile Systeme	→
Infektion mit Schadsoftware über Internet und Intranet	↑
Menschliches Fehlverhalten und Sabotage	→
Kompromittierung von Extranet und Cloud-Komponenten	↗
Social Engineering und Phishing	→
(D)DoS Angriffe	→
Internet-verbundene Steuerungskomponenten	↗
Einbruch über Fernwartungszugänge	↗
Technisches Fehlverhalten und höhere Gewalt	→
Soft- und Hardwareschwachstellen in der Lieferkette	↑

2022 (BSI-CS_005)

Politische und gesetzliche Lage



Politische Lage und gesetzliche Situation

StromVG (2007)

– Art. 8 Aufgaben der Netzbetreiber

[1] Die Netzbetreiber koordinieren ihre Tätigkeiten. Ihnen obliegt insbesondere:

a. die Gewährleistung eines sicheren, leistungsfähigen und effizienten Netzes;
b. die Organisation der Netznutzung und die Regulierung des Netzes unter Berücksichtigung des Austausches mit anderen Netzen;
c. die Bereitstellung der benötigten Reserveleitungskapazität;
d. die Erarbeitung der technischen und betrieblichen Mindestanforderungen für den Netzbetrieb. Sie berücksichtigen dabei internationale Normen und Empfehlungen anerkannter Fachorganisationen.

Politische Lage und gesetzliche Situation

- 2007 StromVG
- 2012 Nationale Strategie zum Schutz der Schweiz vor Cyber-Risiken I
- 2015 SKI-Strategie (Schutz kritischer Infrastruktur) Stromversorgung
- 2018 Revidiertes Energiegesetz
- 2018 Nationale Strategie zum Schutz der Schweiz vor Cyber-Risiken II
- 2018 VSE Handbuch „Grundschutz für Operational Technology (OT)“ in der Stromversorgung
- 2024 …

Regulatorisches Umfeld (Basis BFE)

- Meldepflicht ab 1.1.2025 für Vorfälle in der Informationssicherheit auf Basis des neuen Informationssicherheits-Gesetzes (Rev. ISG)
- Gesetzlich vorgeschriebener IKT-Minimalstandard ab Mitte 2024 im revidierten StromVG. Einhaltung von vorgeschriebenen Profilen (A-D) mit entsprechenden Maturitäten auf Basis Stromlieferung bzw. Stromproduktion voraussichtlich per Mitte 2024.
- Umsetzung mit geplanten Self-Assessments durch die Versorgungs-unternehmen zur Information/Kontrolle an ELCOM

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Finanzdepartement EFD
Nationales Zentrum für Cybersicherheit NCSC

Warum und wann wird einen Meldepflicht eingeführt?

Ziele der Meldepflicht

- **Frühwarnung und Übersicht zur Bedrohungslage:** mehr Informationen über Cyberangriffe ermöglichen es dem NCSC andere Organisationen schneller und präziser zu Warnen und eine gute Übersicht zur Bedrohungslage zu erhalten.
- **Rechtssicherheit- und gleichheit:** der freiwillige Informationsaustausch war lange sehr effizient. Er führt allerdings zum Problem des «Freeriding». Alle profitieren von den geteilten Informationen aber nicht alle sind bereit dazu, Informationen über Cyberangriffe zu teilen.
- **Internationaler Kontext:** mit der NIS-Direktive hat die EU 2018 eine Meldepflicht für Cyberangriffe für alle Mitgliedsstaaten eingeführt.

Die Meldepflicht im Informationssicherheitsgesetz (ISG)

- Das ISG ist ein relativ neues Gesetz (Beschluss 18. Dezember 2020), welches bisher ausschliesslich die Informationssicherheit des Bundes und teilweise der Kantone regelt.
- Das ISG tritt per 1. Januar 2024 in Kraft.
- Die Einführung der Meldepflicht wird als Revision des ISG umgesetzt. Das ISG wird so erweitert zu einem Informationssicherheitsgesetz mit Auswirkungen auf kritische Infrastrukturen.

➔ Die Meldepflicht tritt noch nicht am 1. Januar 24 in Kraft. Sie wird separat als Revision des Gesetzes beschlossen.

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Finanzdepartement EFD
Nationales Zentrum für Cybersicherheit NCSC

Wer muss was auf welche Art und Weise melden?

WER muss melden (Art. 74*b*) – Betreiberinnen kritischer Infrastrukturen

- Grundgedanke: aufgeführt werden jene in der Strategie zum Schutz kritischer Infrastrukturen aufgelisteten Teilsektoren, welche gegenüber Cyberangriffen verwundbar sind.
- Betroffen sind insgesamt 19 Bereiche
- Für die Definition der Adressaten wird auf bestehende Gesetze verwiesen.
- Art. 74*c:* Der Bundesrat kann die Meldepflicht durch geeignete Kriterien in den jeweiligen Sektoren einschränken, wenn:
  - geringen Abhängigkeit von Informatikmitteln besteht
  - Ausfälle oder Funktionsstörungen der Infrastruktur nur geringe Auswirkungen hätten (Anzahl Personen, Substituierbarkeit, geringe volkswirtschaftliche Bedeutung)

Welche Angriffe müssen gemeldet werden (Art. 74d)?

Ein Cyberangriff muss gemeldet werden, wenn er:

a. die Funktionsfähigkeit der betroffenen kritischen Infrastruktur gefährdet;

b. zu einer Manipulation oder zu einem Abfluss von Informationen geführt hat;

c. über einen längeren Zeitraum unentdeckt blieb, insbesondere wenn Anzeichen dafür bestehen, dass er zur Vorbereitung weiterer Cyberangriffe ausgeführt wurde; oder

d. mit Erpressung, Drohung oder Nötigung verbunden ist.

Inhalt und Frist der Meldung (Art. 74e)

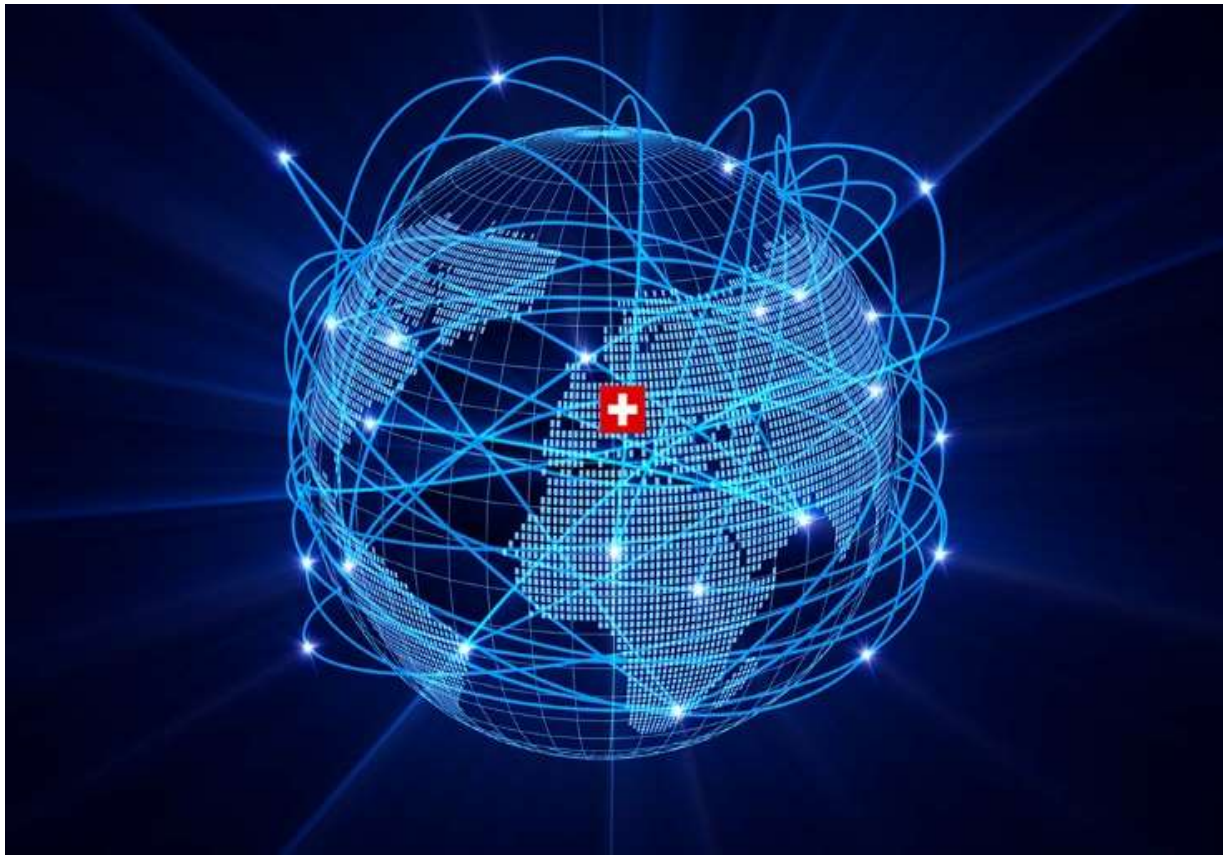
- Die Meldung muss **innert 24 Stunden** nach der Entdeckung des Cyberangriffs erfolgen.
- Sie muss Informationen zur meldepflichtigen Behörde oder Organisation, zur Art und Ausführung des Cyberangriffs, zu seinen Auswirkungen, zu ergriffenen Massnahmen und, soweit bekannt, zum geplanten weiteren Vorgehen enthalten.
- Sind zum Zeitpunkt der Meldung nicht alle erforderlichen Informationen bekannt, so ergänzt die meldepflichtige Behörde oder Organisation die Meldung, sobald sie über neue Informationen verfügt

Sanktionen

Mehrstufiges Verfahren:

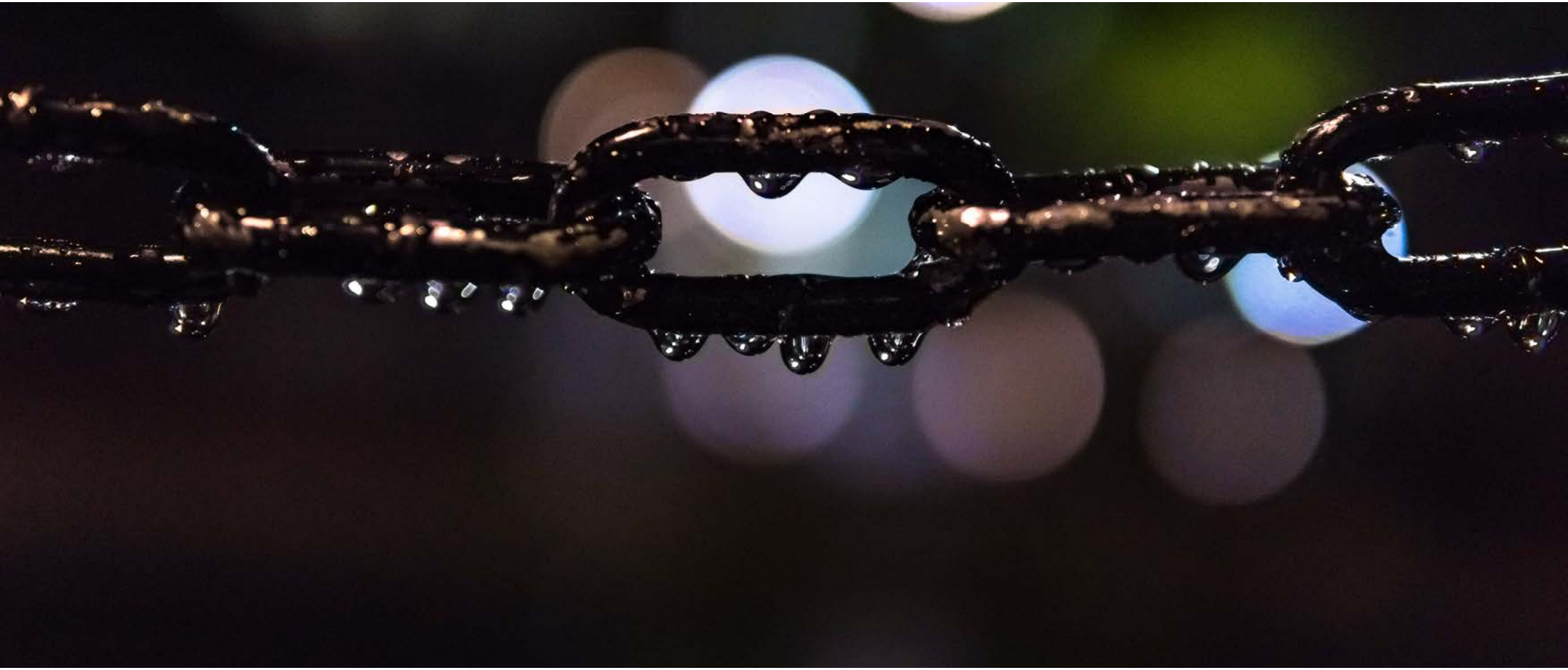
- Das NCSC muss die kritische Infrastruktur auf die Unterlassung aufmerksam machen.
- Kommt die Betreiberin trotz dieser Information ihrer Pflicht nicht nach, so erlässt das NCSC eine Verfügung über die umzusetzenden Pflichten.
- Wird die Verfügung ignoriert, erstattet das NCSC Strafanzeige. Möglich sind Bussen bis CHF 100’000.

Vielen Dank für die Aufmerksamkeit

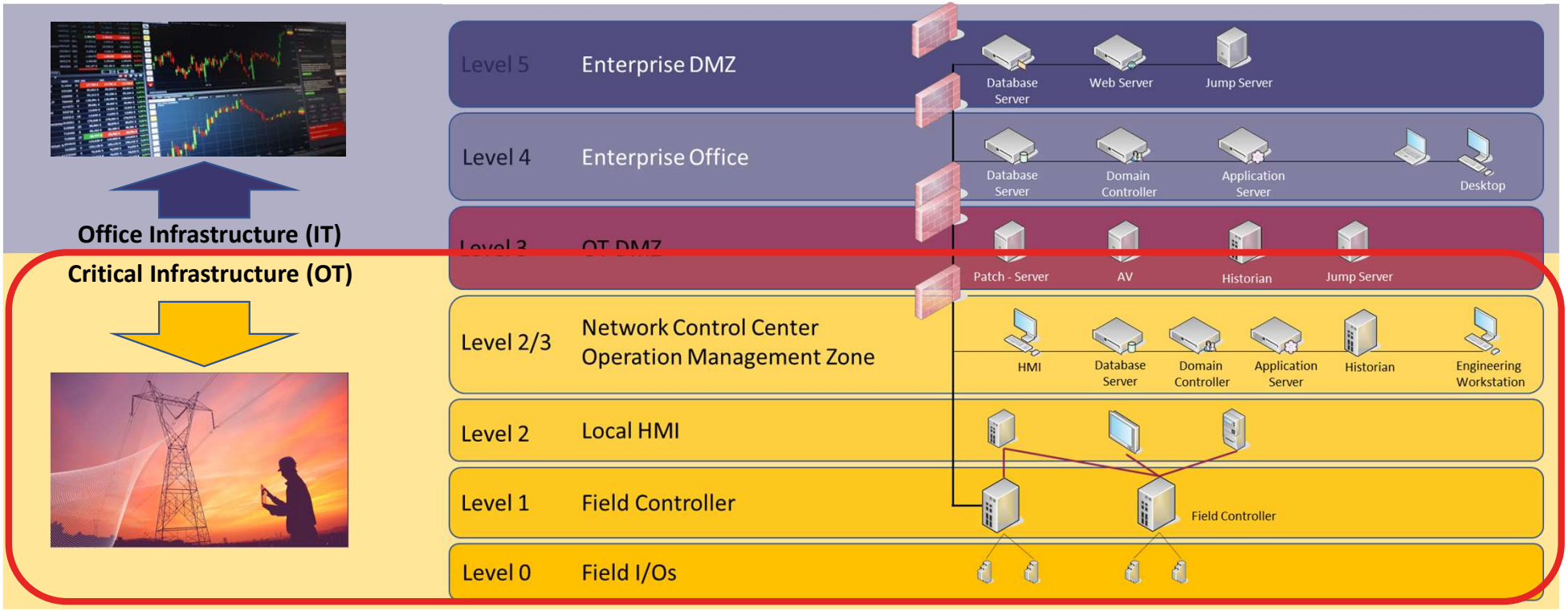


Manuel Suter
Leiter Geschäftsstelle
NCSC

IT und…oder…ohne…mit…OT….



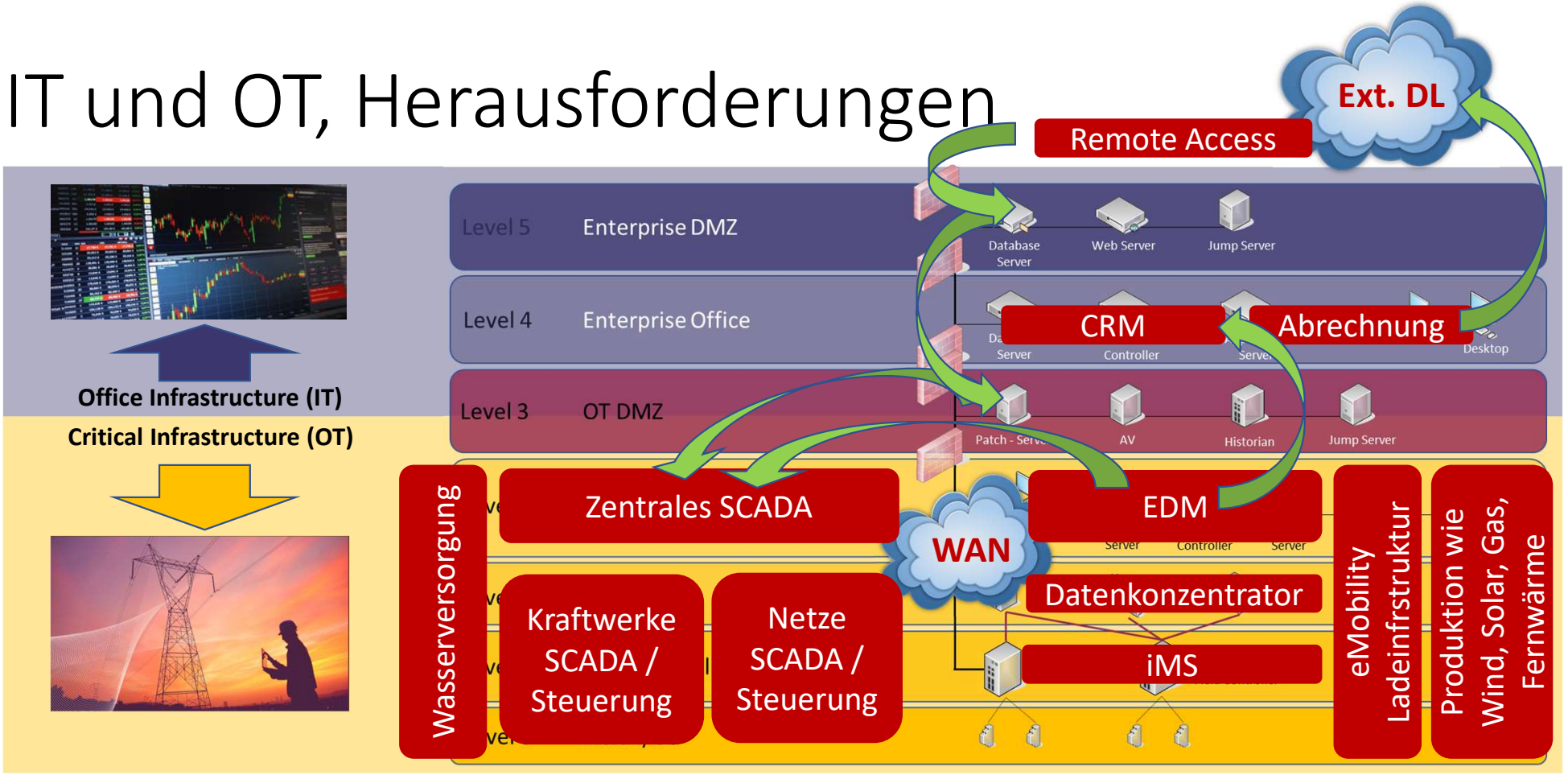
IT und OT, was ist der Unterschied?



IT und OT, was ist der Unterschied?

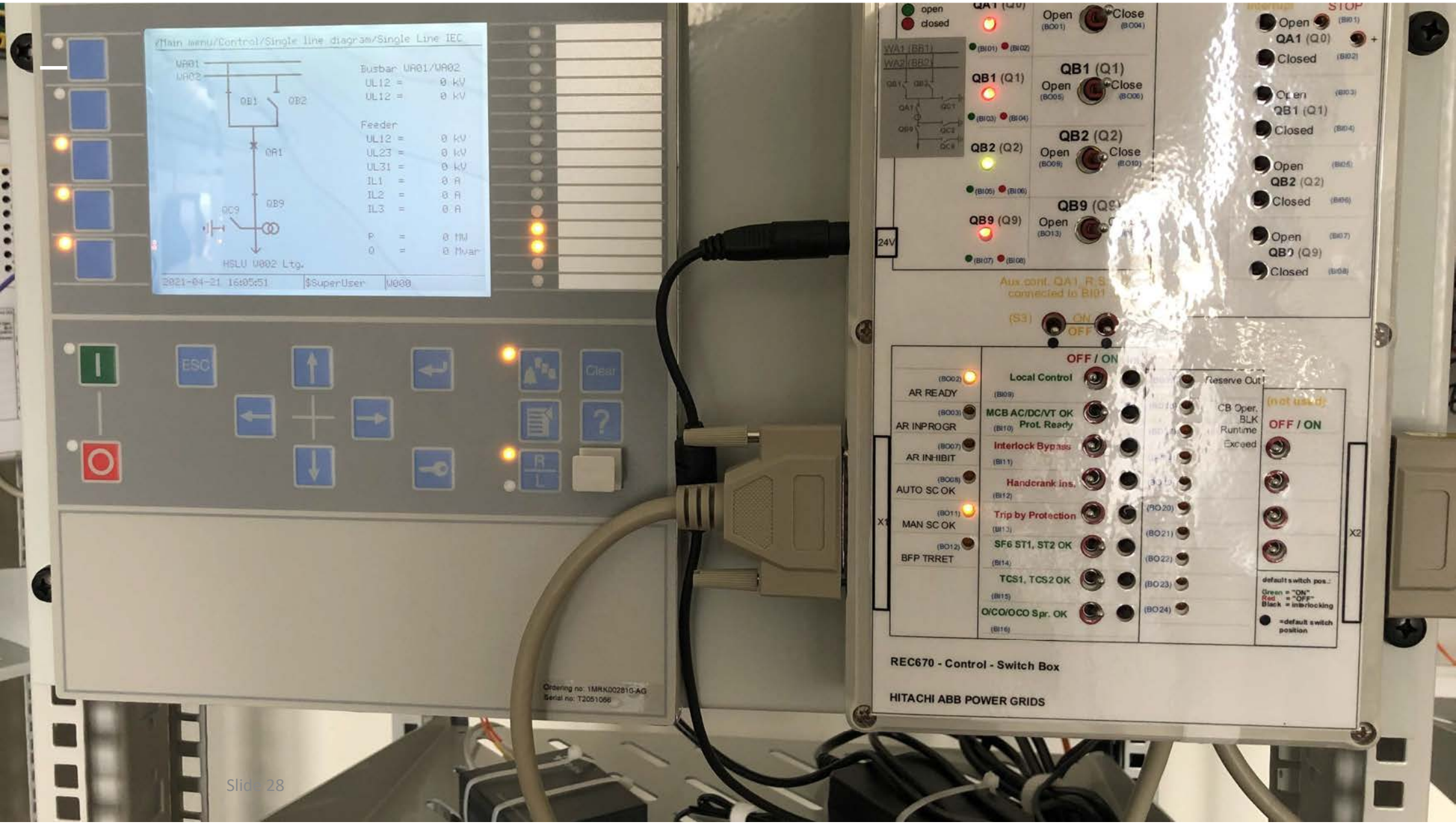
Common IT	OT
Standard IP Protokolle	Non-Standard Protokolle (Bus, Verkabelung etc.)
Aktuelle OS (Windows, Linux etc.)	Ältere / Embeded OS
Wartungsfenster für reguläres Patching/Upgrade	Ausserbetriebnahmen für Patching/Upgrade
Aktives Scanning	Passives Scanning, ansonsten Störung der Systeme
Life Cycle 5 Jahre	Life Cycle 15 Jahre und mehr
Generelles IT Wissen	Spezifisches Wissen
M2H (Mensch ist auch ein Sensor)	M2M (Maschine kommuniziert mit Maschine)

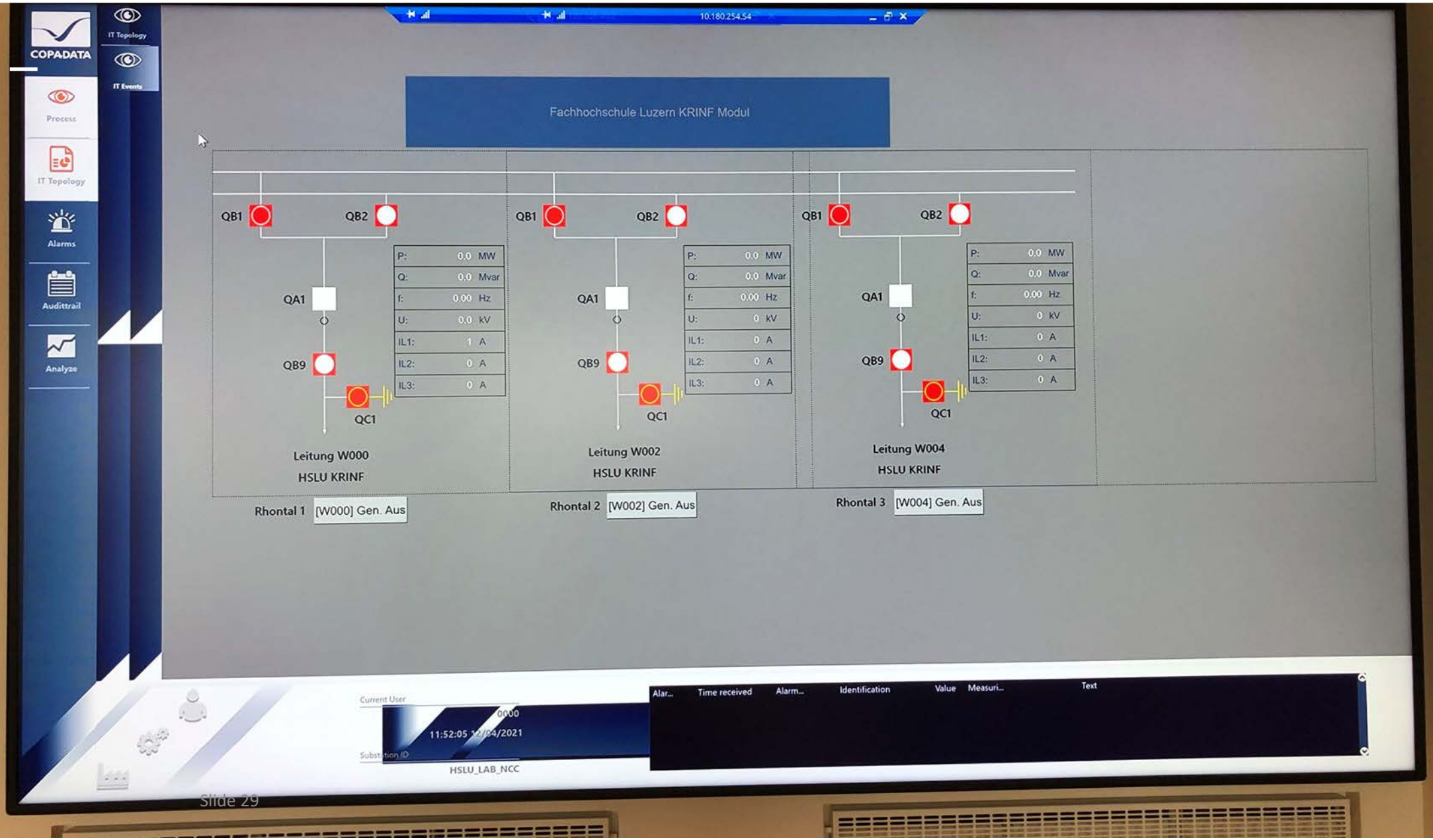
IT und OT, Herausforderungen



KRINFLAB

27 22.11.2023


/Main menu/Control/Single line diagram/Single Line IEC
WA01
WA02
QB1
QB2
QA1
QC9
QB9
HSLU U002 Ltg.
Busbar WA01/WA02
UL12 = 0 kV
UL12 = 0 kV
Feeder
UL12 = 0 kV
UL23 = 0 kV
UL31 = 0 kV
IL1 = 0 A
IL2 = 0 A
IL3 = 0 A
P = 0 MW
Q = 0 Mvar
2021-04-21 16:05:51
$SuperUser
U000
ESC
Clear
Ordering no: 1MRK002810-AG
Serial no: T2051056
open
closed
QB1 (Q1)
QB2 (Q2)
QB9 (Q9)
Open
Close
24V
OFF / ON
Local Control
MCB AC/DC/VT OK
Prot. Ready
Interlock Bypass
Handcrank ins.
Trip by Protection
SF6 ST1, ST2 OK
TCS1, TCS2 OK
O/CO/OCO Spr. OK
AR READY
AR INPROGR
AR INHIBIT
AUTO SC OK
MAN SC OK
BFP TRRET
Reserve Out
CB Oper. BLK
Runtime Exceed
(not used)
OFF / ON
default switch pos.:
Green = "ON"
Red = "OFF"
Black = interlocking
=default switch position
REC670 - Control - Switch Box
HITACHI ABB POWER GRIDS

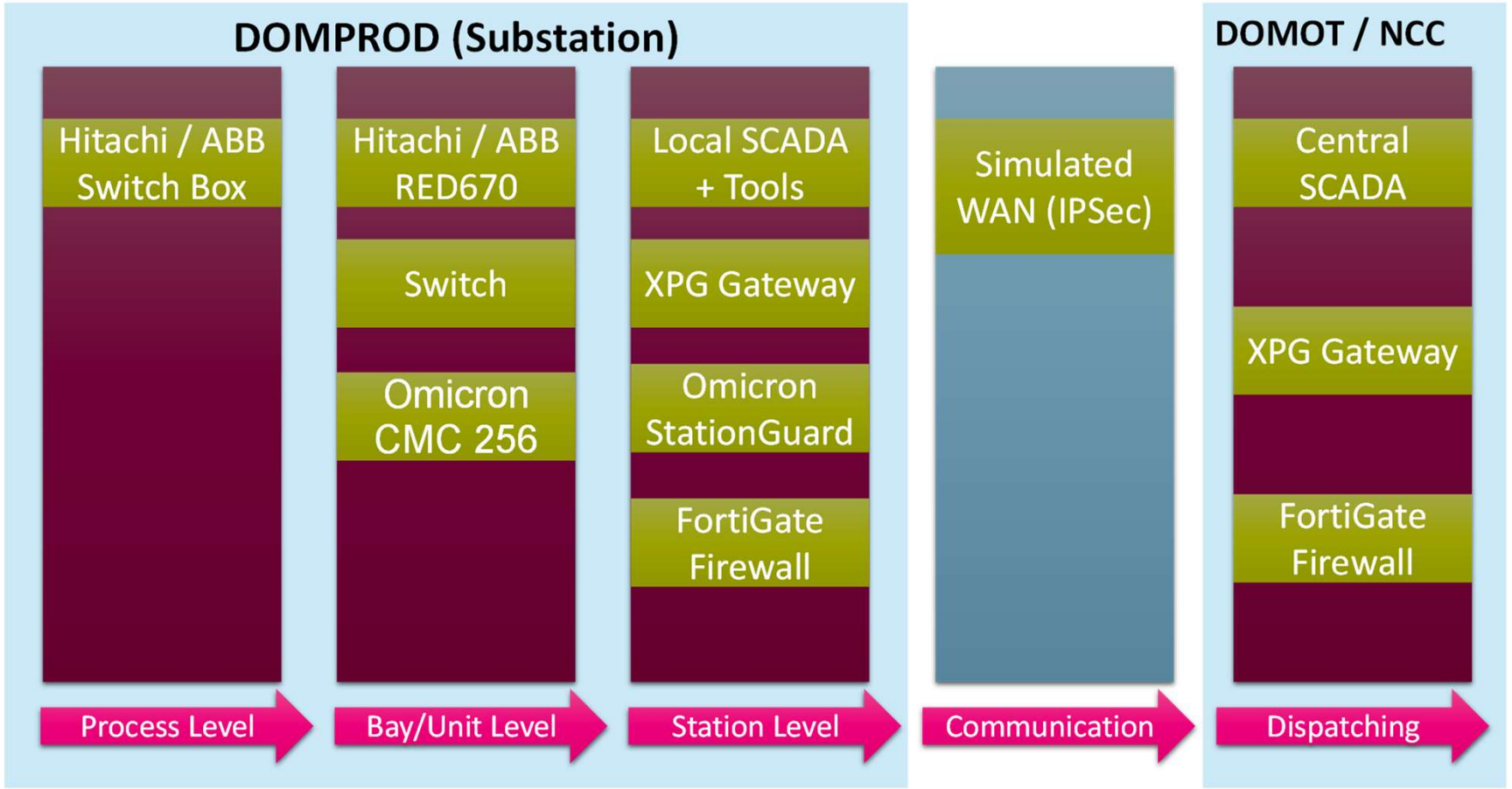

COPADATA
IT Topology
IT Events
Process
IT Topology
Alarms
Audittrail
Analyze
10.180.254.54
Fachhochschule Luzern KRINF Modul
QB1
QB2
QA1
QB9
QC1
P: 0.0 MW
Q: 0.0 Mvar
f: 0.00 Hz
U: 0.0 kV
IL1: 1 A
IL2: 0 A
IL3: 0 A
Leitung W000
HSLU KRINF
Rhontal 1
[W000] Gen. Aus
QB1
QB2
QA1
QB9
QC1
P: 0.0 MW
Q: 0.0 Mvar
f: 0.00 Hz
U: 0 kV
IL1: 0 A
IL2: 0 A
IL3: 0 A
Leitung W002
HSLU KRINF
Rhontal 2
[W002] Gen. Aus
QB1
QB2
QA1
QB9
QC1
P: 0.0 MW
Q: 0.0 Mvar
f: 0.00 Hz
U: 0 kV
IL1: 0 A
IL2: 0 A
IL3: 0 A
Leitung W004
HSLU KRINF
Rhontal 3
[W004] Gen. Aus
Current User
0000
11:52:05
Substation ID
HSLU_LAB_NCC
Alar...
Time received
Alarm...
Identification
Value
Measuri...
Text

Aufbau der Laborumgebung

„Entspricht einem typischen Energie-KMU“ - Firma Alsec

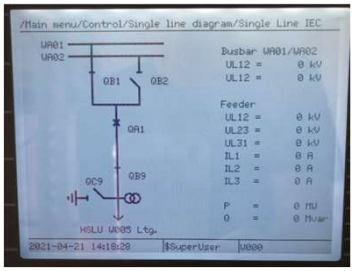
6x Unterwerk (Substation)

1x Leitstelle (Control Center)

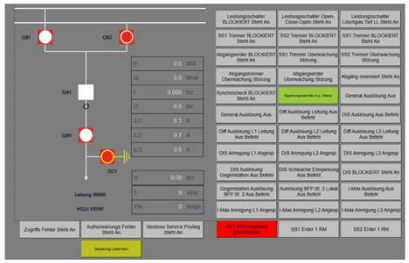


Labor-Übungen

Schutzfunktionen prüfen



SCADA System bedienen

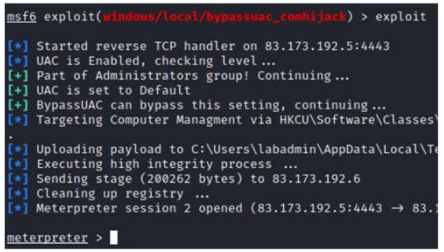


Cyber Maturity Assessment

IDS Integration

Gateway absichern

Cyber Angriffe durchführen



„Building the bridge“ zwischen IT und OT

Fragen und Antworten?

Wir stehen für eine SICHERE Energieversorgung!

Danke für Ihre Aufmerksamkeit!

www.alsec.ch

HSLU Hochschule Luzern

Hinweis

Diese Dokumente sind ausschliesslich für die Teilnehmer dieser Veranstaltung bestimmt. Aus urheberrechtlichen Gründen wird keine andere Verbreitung der Dokumente oder von Auszügen erlaubt. Die Urheberrechte verbleiben dabei beim jeweiligen Autor.

oneconsult
together against cyberattacks

Pentesters Diary

Geschichten aus dem Alltag eines Penetration Testers

HSLU, Information & Cyber Security Tech-Meeting

Yves Kraft | HSLU Rotkreuz, 23. November 2023

Yves Kraft

Branch Manager Bern

Head of Cyber Security Academy

Pentesters Diary
11 Geschichten aus der alltäglichen Arbeit als Pentester

- (1) Handies sind Energieverschwender
- (2) C:\Temp
- (3) Shutdownmanyservers.bat
- (4) Fehlermeldung
- (5) OT-Audit: ”Naja, dann ist immerhin wieder mal abgestaubt”
- (6) SSTI
- (7) Die Schweizerische Post erhöht die Preise für Porto
- (8) GeilerTyp1!
- (9) Have I been Pwned?
- (10) 0000
- (11) Darf ich etwas zeichnen?

Let’s connect

www.oneconsult.com

/oneconsult-ag

/OneconsultAG

/oneconsult

Kontakt

Holding

Oneconsult International AG
Giesshübelstrasse 45
8045 Zürich
Schweiz

+41 43 377 22 22
info@oneconsult.com

Schweiz

Oneconsult AG
Giesshübelstrasse 45
8045 Zürich
Schweiz

+41 43 377 22 22
info@oneconsult.com

Oneconsult AG
Aarbergergasse 56
3011 Bern
Schweiz

+41 31 327 15 15
info@oneconsult.com

Deutschland

Oneconsult Deutschland AG
Agnes-Pockels-Bogen 1
80992 München
Deutschland

+49 89 248820 600
info@oneconsult.com

Neuseeland

Oneconsult New Zealand Limited
Level 3, 33-45 Hurstmere Road
Takapuna, Auckland 0622
New Zealand

+64 27 325 4299
info@oneconsult.com

Cybersecurity in the Era of AI

Information & Cyber Security Tech-Meeting
22.11.2023

Nico Heise
Security Architect
Nico.Heise@ibm.com

Agenda

- Introduction - The (re)rise of AI
- Security for AI
- AI for Security

Artificial Intelligence (AI)

Human intelligence exhibited by machines

Learning, reasoning, perceiving, and problem solving.

Machine Learning (ML)

Systems that learn from historical data

Discover patterns and generate corresponding outputs

Deep Learning (DL)

ML technique that mimics human brain function

Enable complex applications, like image and speech recognition.

Foundation Model

Generative AI systems

Generate sequences of related data elements (for example, like a sentence).

Security and AI

Security for AI

Secure the underlying AI training data

Secure model development

Secure the usage of AI models

AI for Security

AI will manage repetitive security tasks

AI will generate security content

AI will learn and create active responses

Attacker’s Use of AI

AI Powered Attacks

Generate: DeepHack tool learned SQL injection

Automate: Generate targeted phishing attacks on Twitter

Refine: Neural network powered password crackers

Evade: Generative adversarial networks learn novel steganographic channels

Attacking AI

Poison: Microsoft Tay chatbot poisoning via Twitter (and Watson Urban Dictionary “poisoning”)

Evade: Real-world attacks on computer vision for facial recognition biometrics and autonomous vehicles

Harden: Genetic algorithms and reinforcement learning (OpenAI Gym) to evade malware detectors

Theft of AI

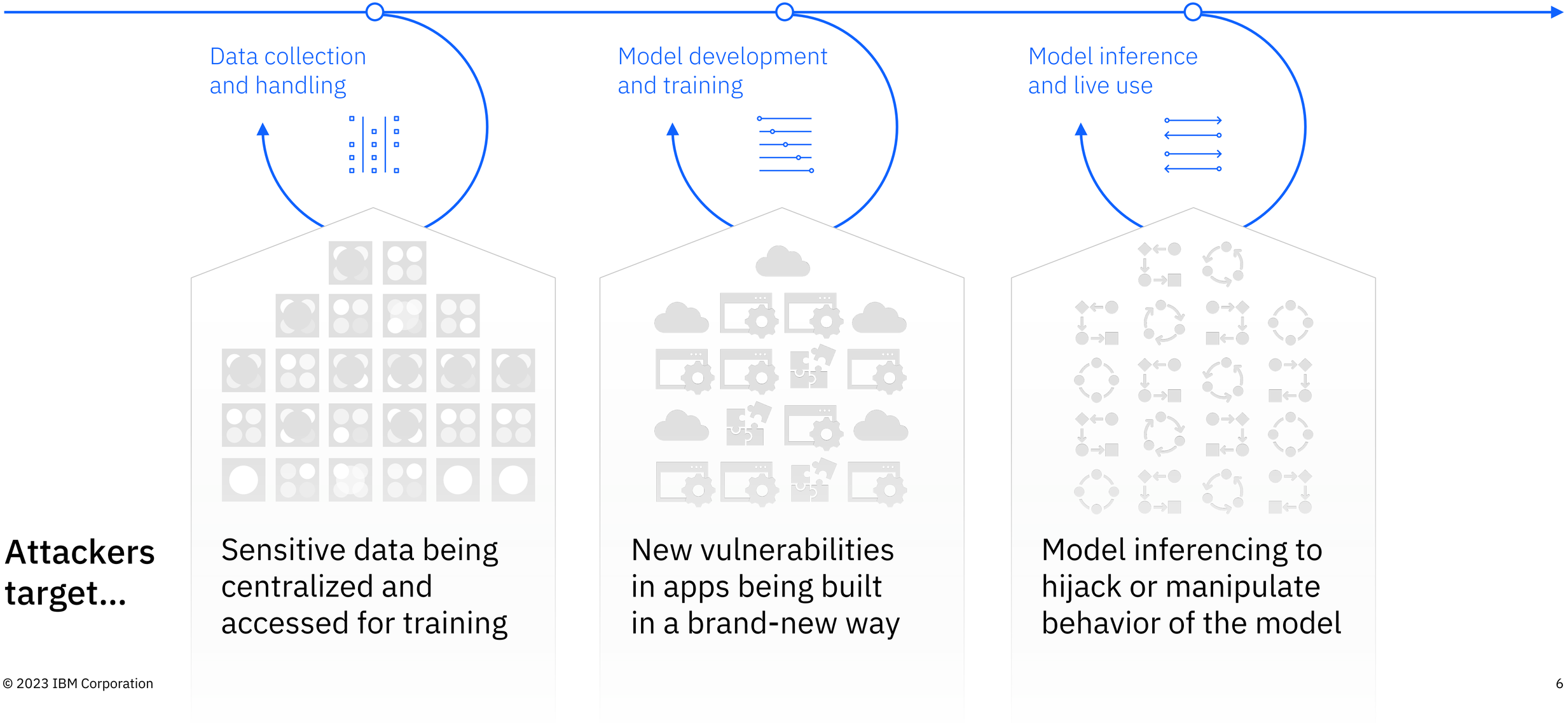
Theft: Stealing machine learning models via public APIs

Transferability: Practical black-box attacks learn surrogate models for transfer attacks

Privacy: Model inversion attacks steal training data

Security for AI - Adversarial risks across the AI pipeline

AI pipeline



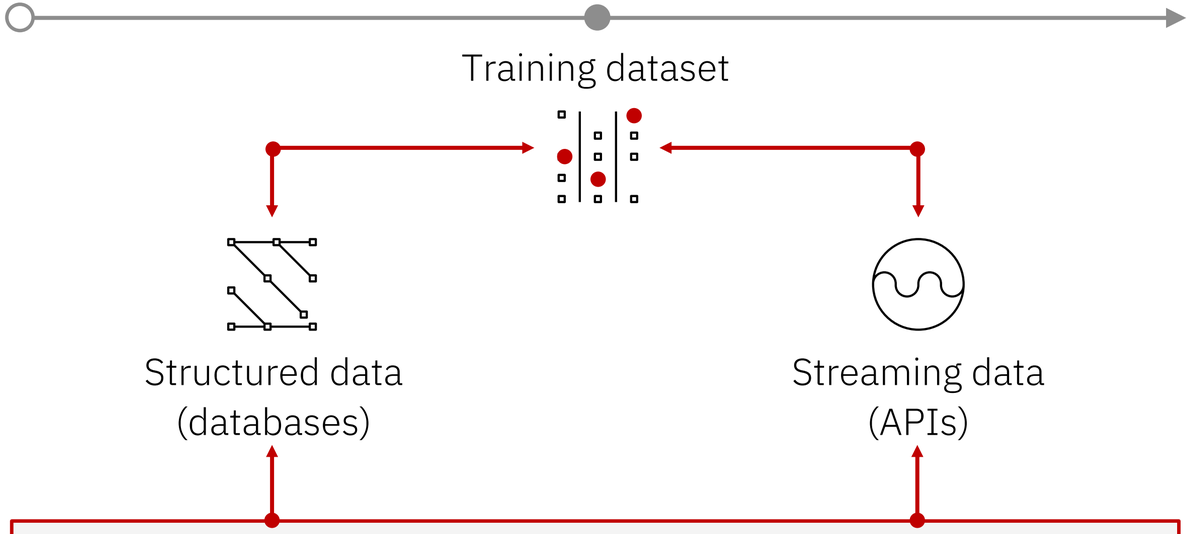
Attackers target...

Sensitive data being centralized and accessed for training

New vulnerabilities in apps being built in a brand-new way

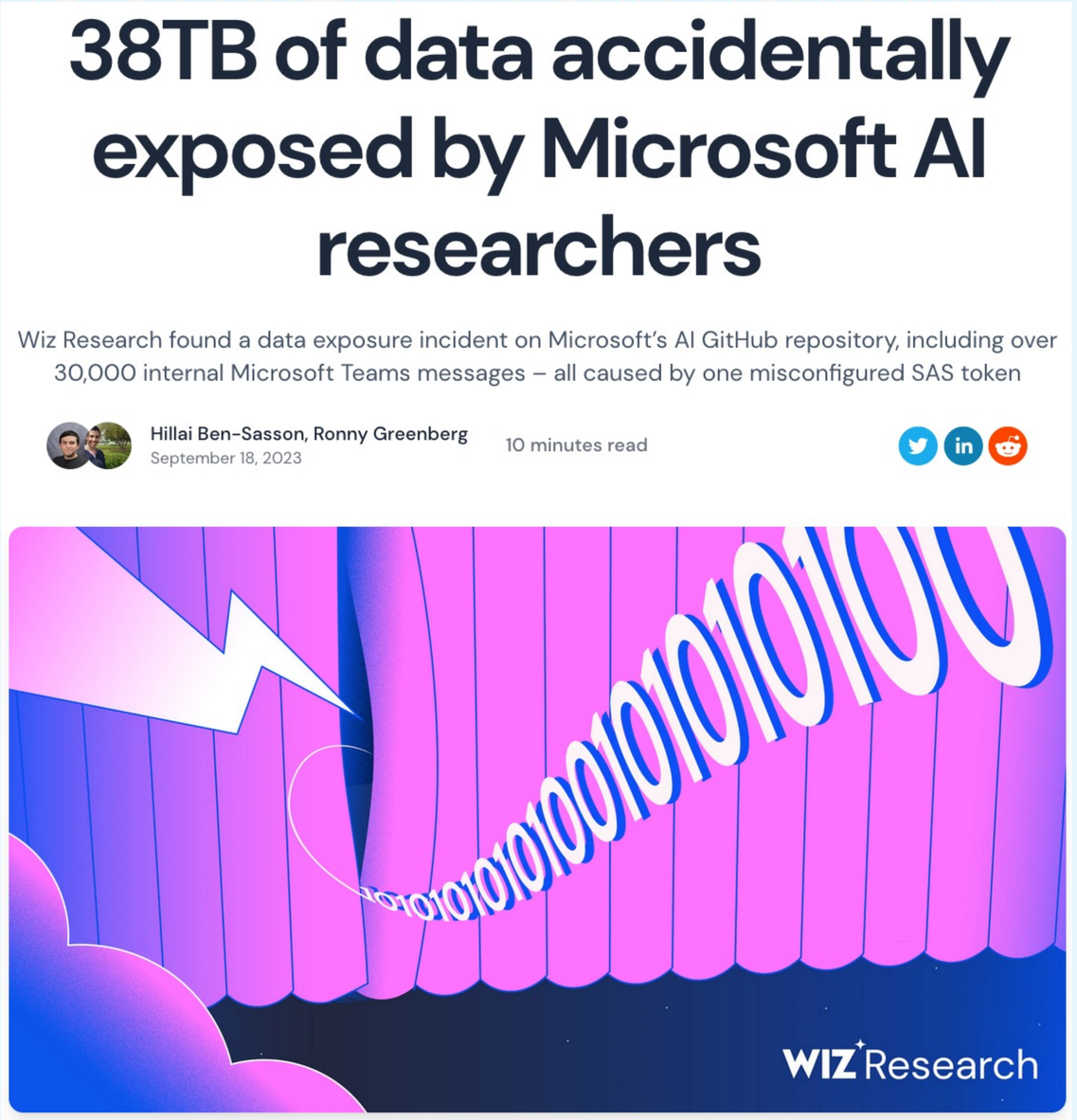
Model inferencing to hijack or manipulate behavior of the model

Data collection and handling risks



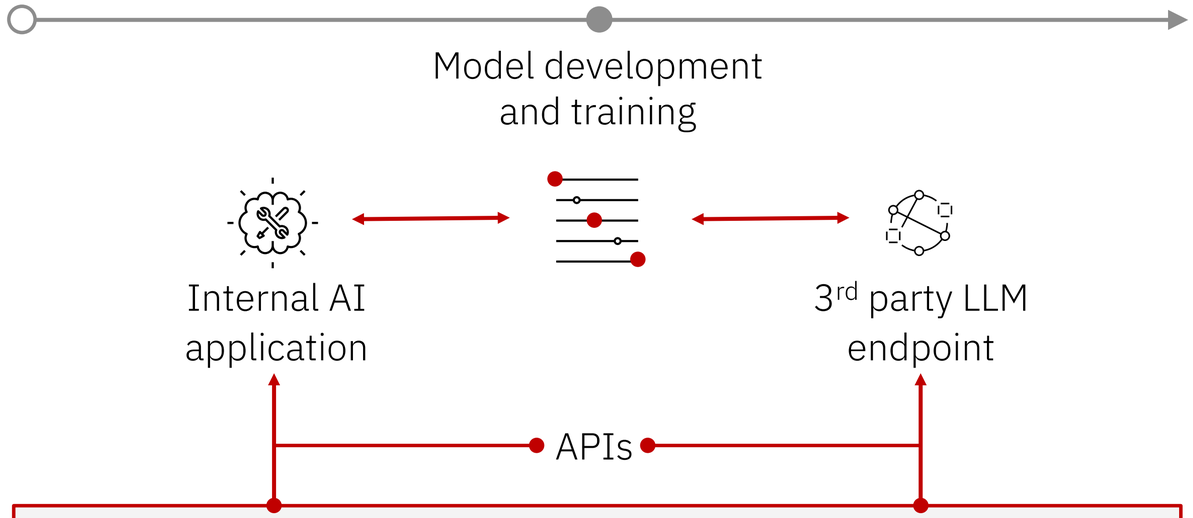
Attackers target the underlying datasets

- Data exfiltration
- Data poisoning



Source: https://www.wiz.io/blog/38-terabytes-of-private-data-accidentally-exposed-by-microsoft-ai-researchers

Model development and training risks



Attackers exploit vulnerabilities and dependencies

- Supply chain attacks
- API attacks
- Privilege escalation

DARKREADING | The Edge | DR Tech | Sections | Events

Machine Learning Models: A Dangerous New Attack Vector

Threat actors can weaponize code within AI technology to gain initial network access, move laterally, deploy malware, steal data, or even poison an organization's supply chain.

Elizabeth Montalbano
Contributor, Dark Reading

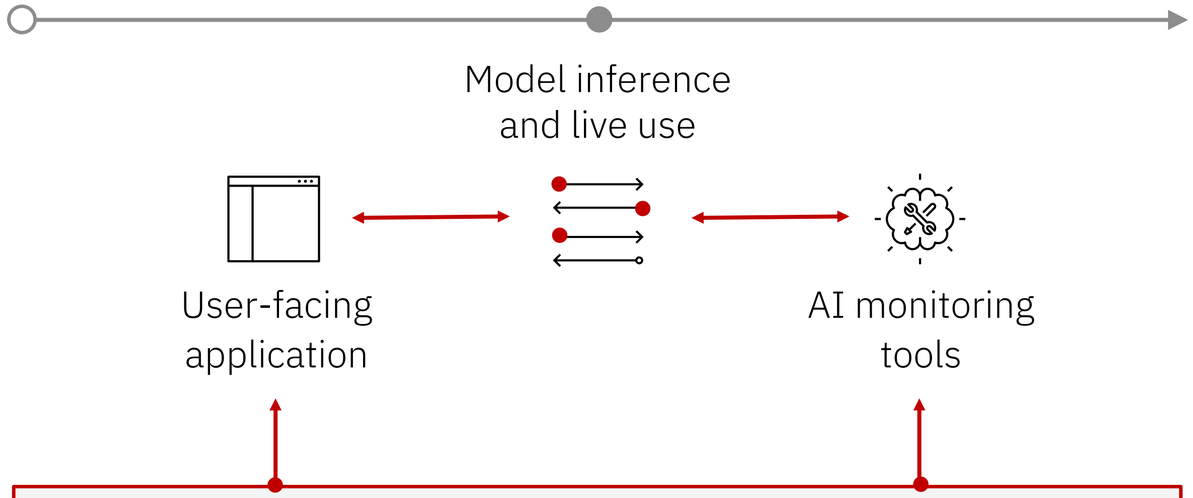
December 06, 2022

Source: Skorzewiak via Alamy Stock Photo

Threat actors can hijack machine learning (ML) models that power artificial intelligence (AI) to deploy malware and move laterally across enterprise networks, researchers have found. These models, which often are publicly available, serve as a new launchpad for a range of attacks that also can poison an organization's supply chain — and enterprises need to prepare.

Source: https://www.darkreading.com/threat-intelligence/machine-learning-models-dangerous-new-attack-vector

Model inference and live use risks



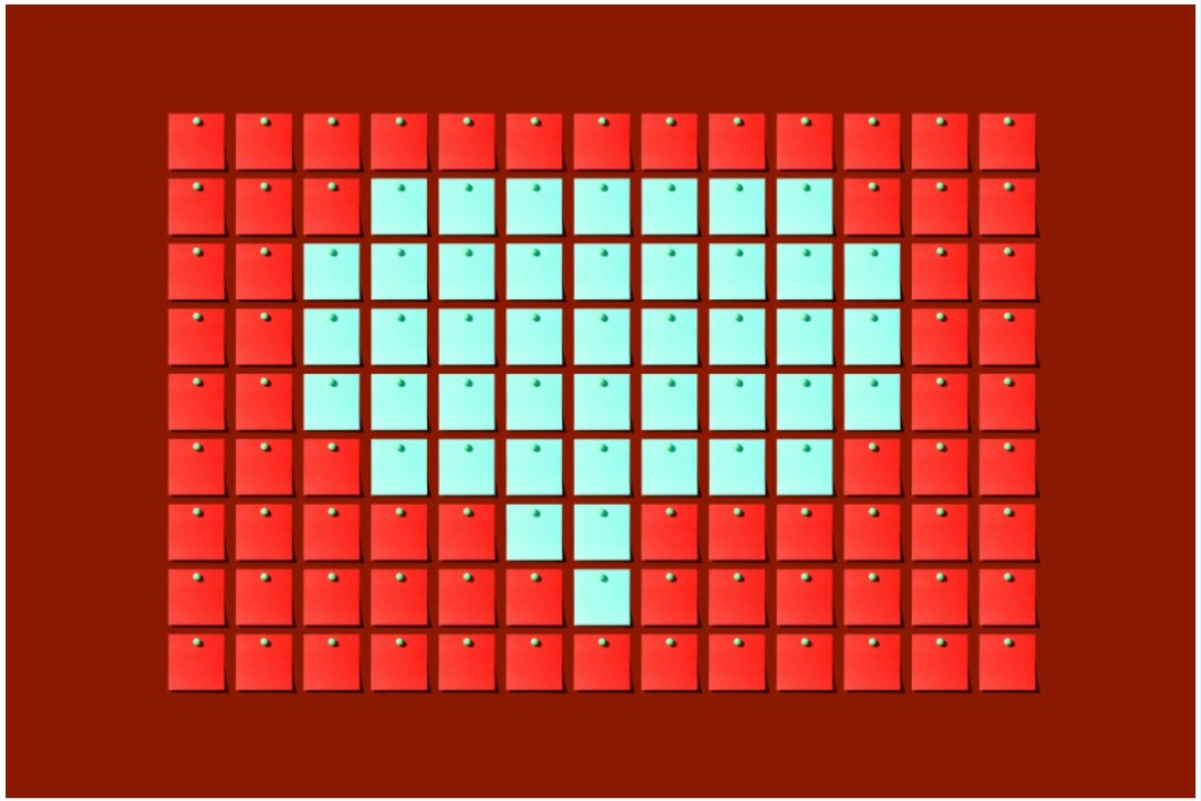
Attackers exploit vulnerabilities and dependencies

- Prompt injection
- Model denial of service
- Model theft

WILL KNIGHT SECURITY AUG 1, 2023 7:00 AM

A New Attack Impacts Major AI Chatbots—and No One Knows How to Stop It

Researchers found a simple way to make ChatGPT, Bard, and other chatbots misbehave, proving that AI is hard to tame.

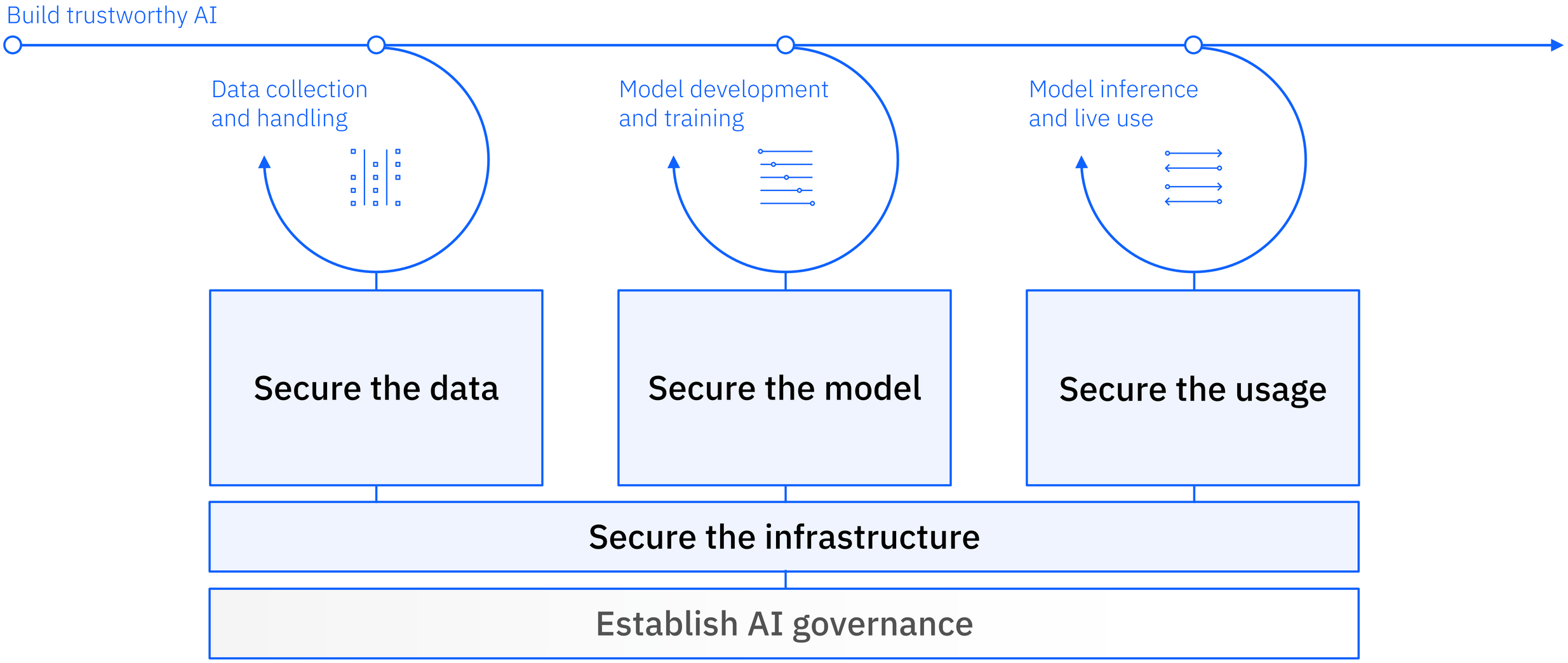


PHOTOGRAPH: MIRAGEC/GETTY IMAGES

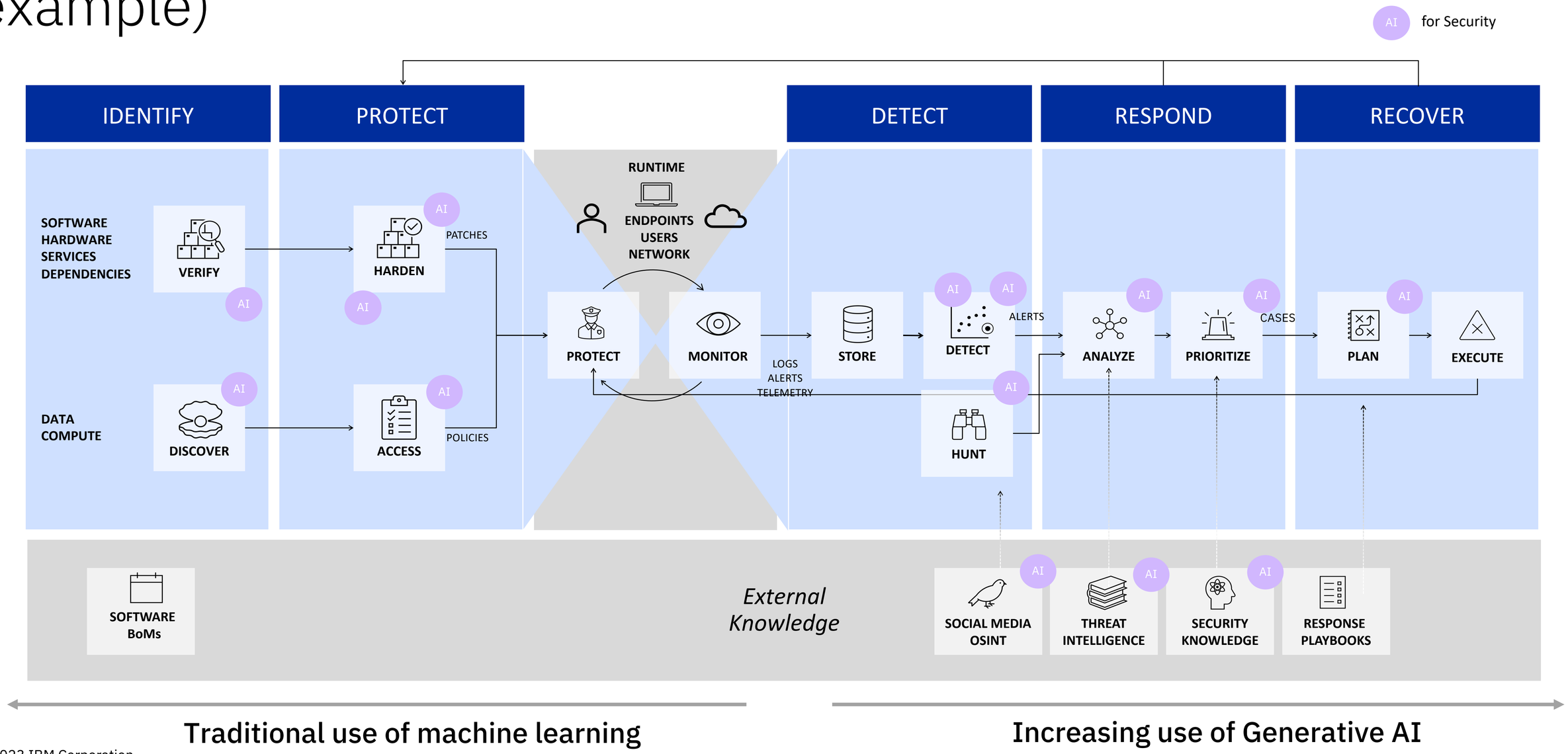
Source: https://www.wired.com/story/ai-adversarial-attacks/

To summarize: What you need to do

Security for AI framework

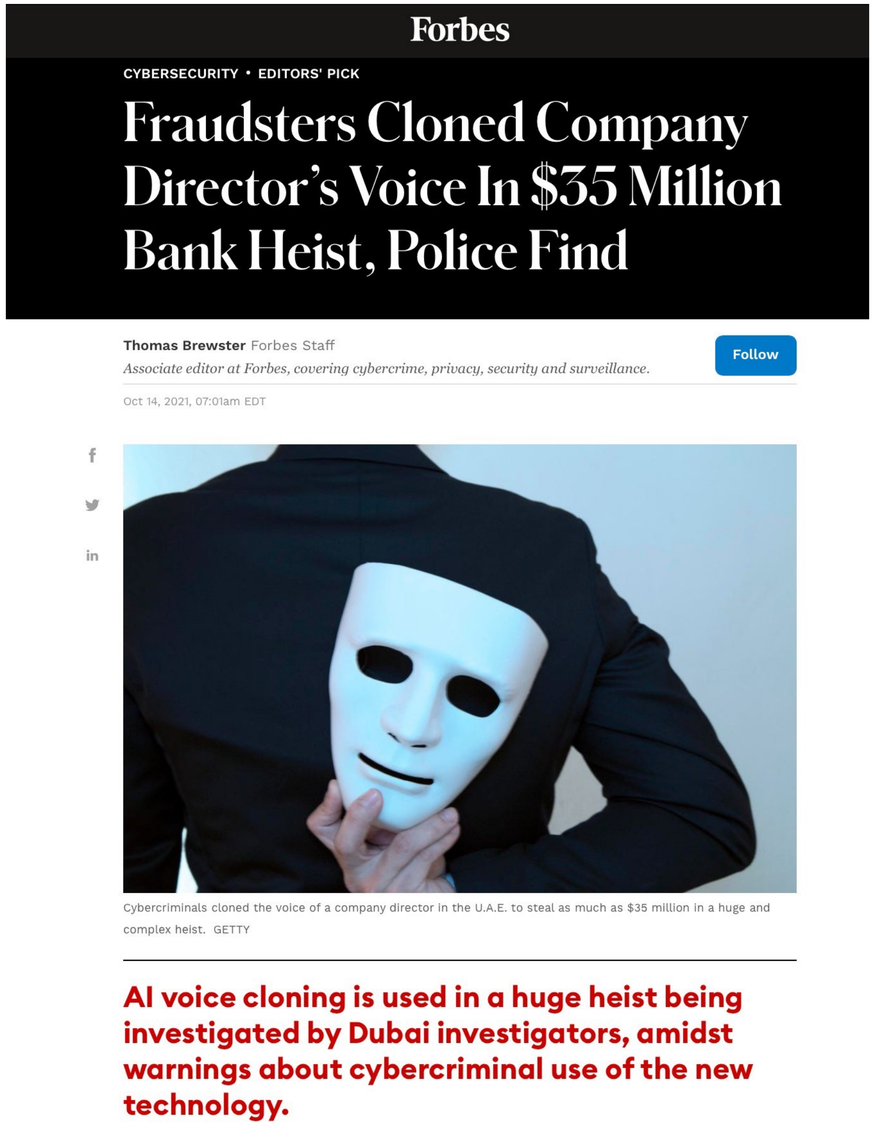


AI for Security - Security Operations (NIST CSF aligned example)



Advanced Use of AI In Social Engineering

- Fraudsters already have leveraged AI to duplicate voice of C-level execs to successfully request wire transfers
- ChatGPT is already pretty good at writing believable phishing emails, despite efforts to limit its ability to do harm
- Chat bots in tandem with AI voice replication will be leveraged to automate the vishing process entirely
- Blackmail may also increase with the continued evolution of deep fake technologies
- Voiceprint authentication solutions will be come obsolete or highly ineffective



Forbes

CYBERSECURITY • EDITORS' PICK

Fraudsters Cloned Company Director's Voice In $35 Million Bank Heist, Police Find

Thomas Brewster Forbes Staff

Associate editor at Forbes, covering cybercrime, privacy, security and surveillance.

Follow

Oct 14, 2021, 07:01am EDT

Cybercriminals cloned the voice of a company director in the U.A.E. to steal as much as $35 million in a huge and complex heist. GETTY

AI voice cloning is used in a huge heist being investigated by Dubai investigators, amidst warnings about cybercriminal use of the new technology.

Source: https://www.forbes.com/sites/thomasbrewster/2021/10/14/huge-bank-fraud-uses-deep-fake-voice-tech-to-steal-millions/

Thank you

Follow us on:

ibm.com/security

securityintelligence.com

ibm.com/security/community

xforce.ibmcloud.com

@ibmsecurity

youtube.com/ibmsecurity