

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

AI: Socio-Legal, Ethical, and Cybersecurity Perspectives

Feb 13, 2025

Prof. Mauro Conti
University of Padua, Italy



SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP



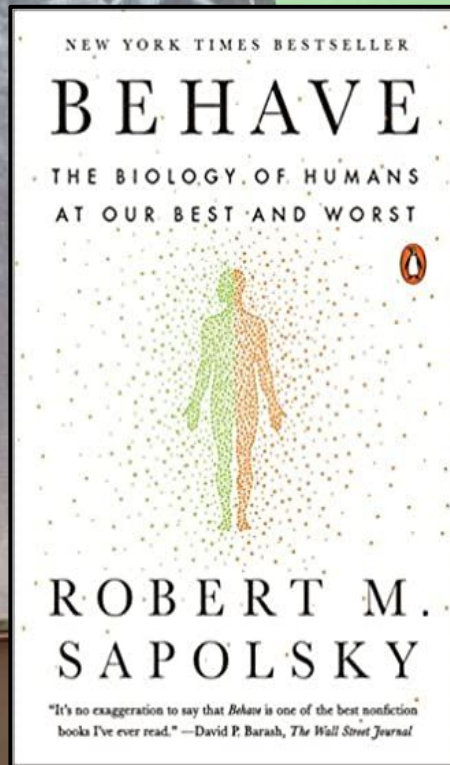
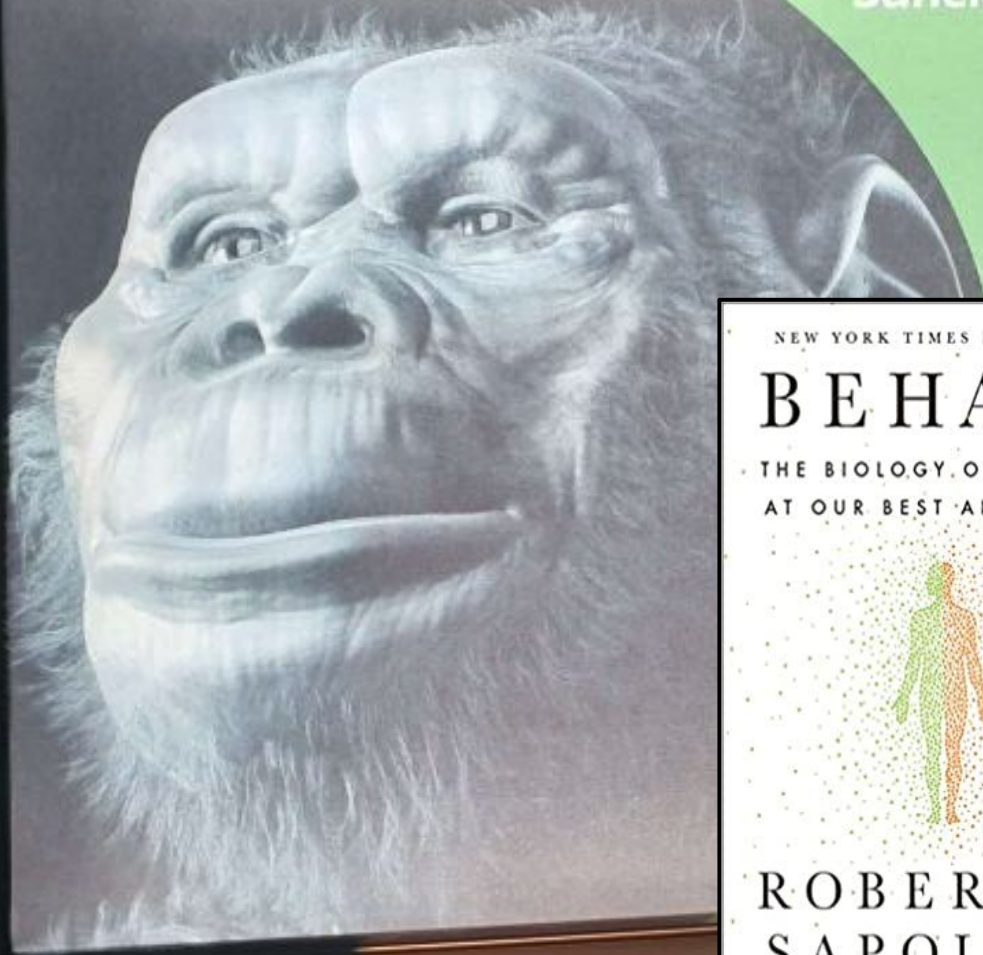
s p r i t z m a t t e r
your cybersecurity partner for innovation
Spin-off Università di Padova



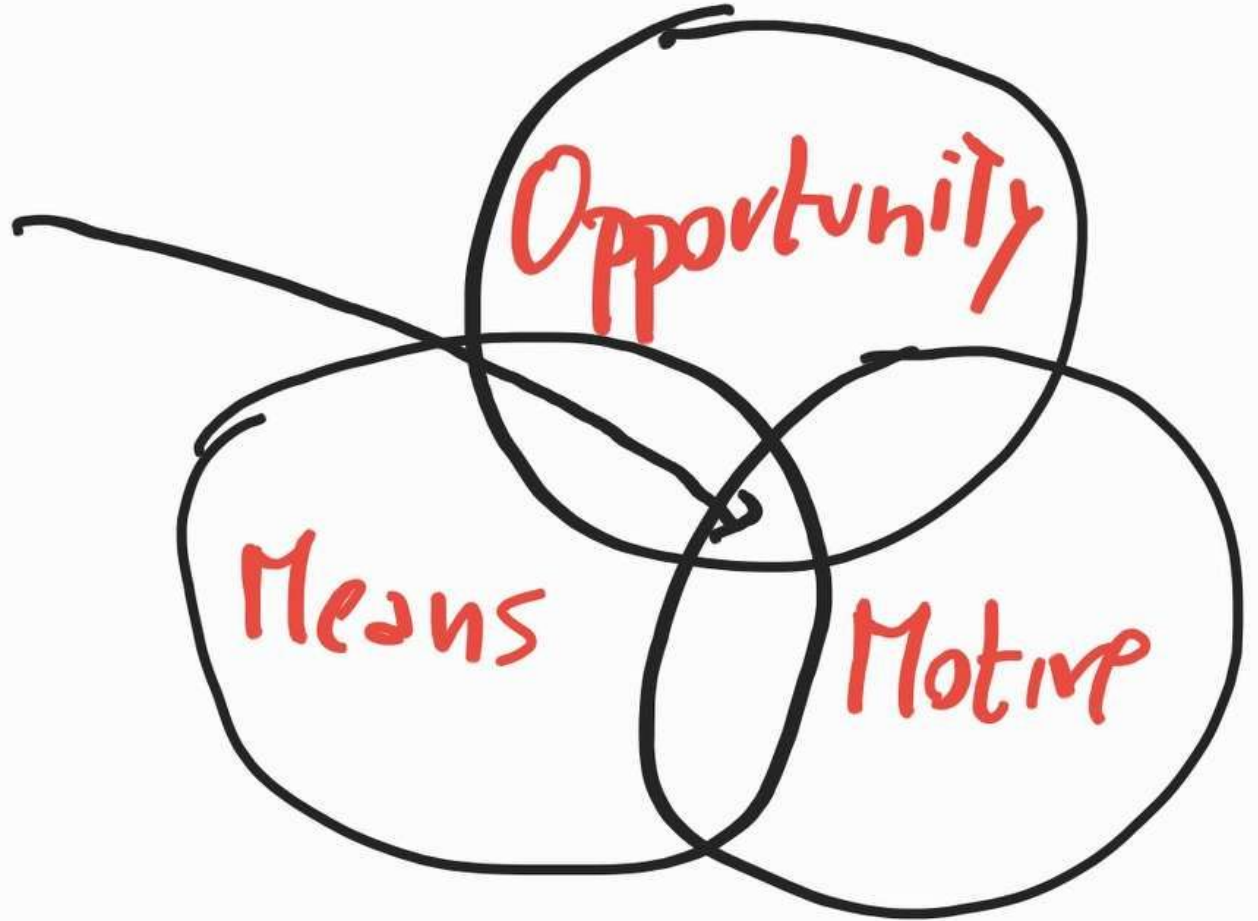


<https://drive.google.com/file/d/1Gch8tWOWcRnw97zeyXOeYR1uZlz6qBKK>

Sahelanthropus tchadensis



Crime





Crime

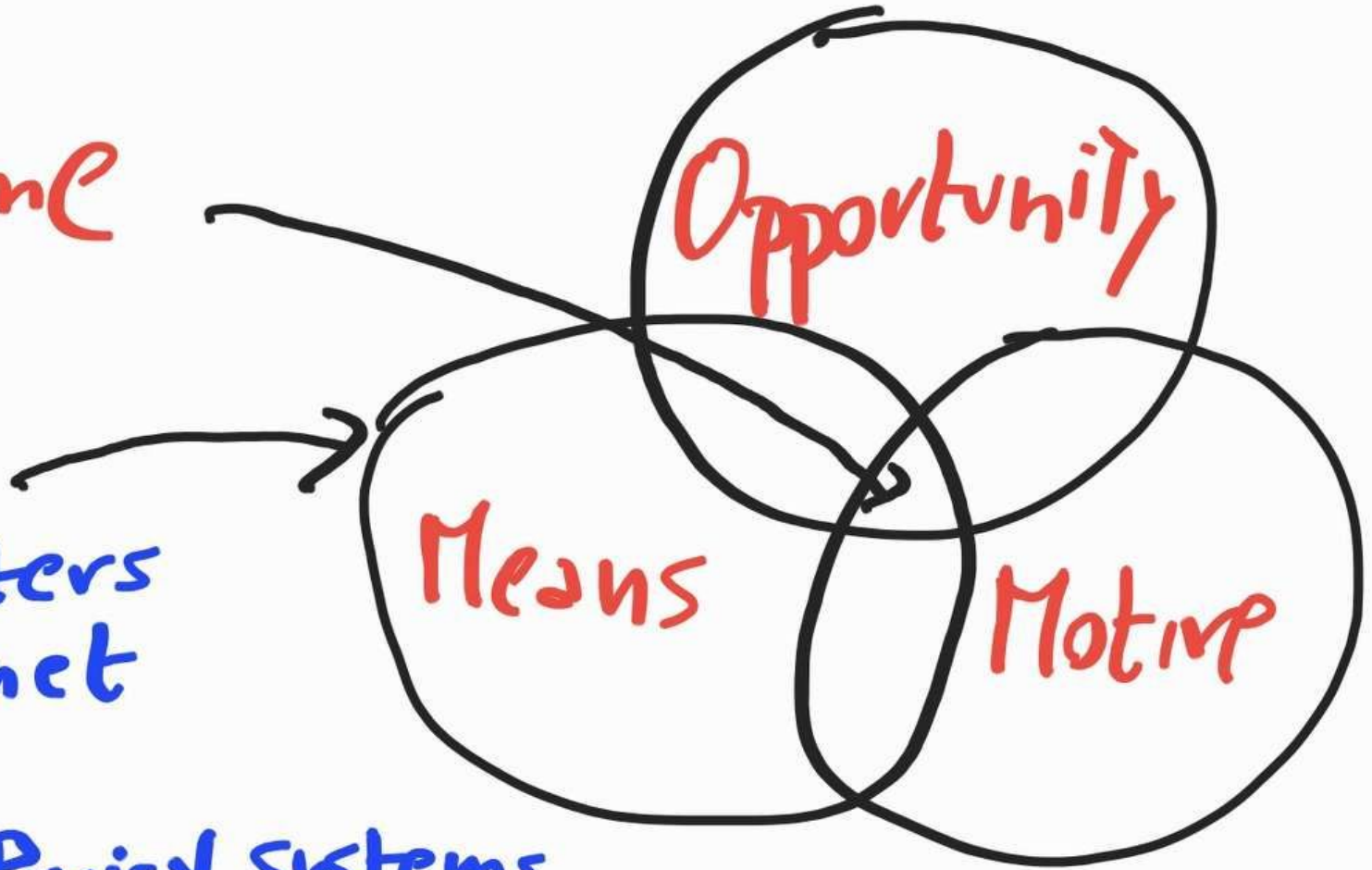
Opportunity

Means

Motive

Computers
Internet
IoT

Cyber Physical Systems

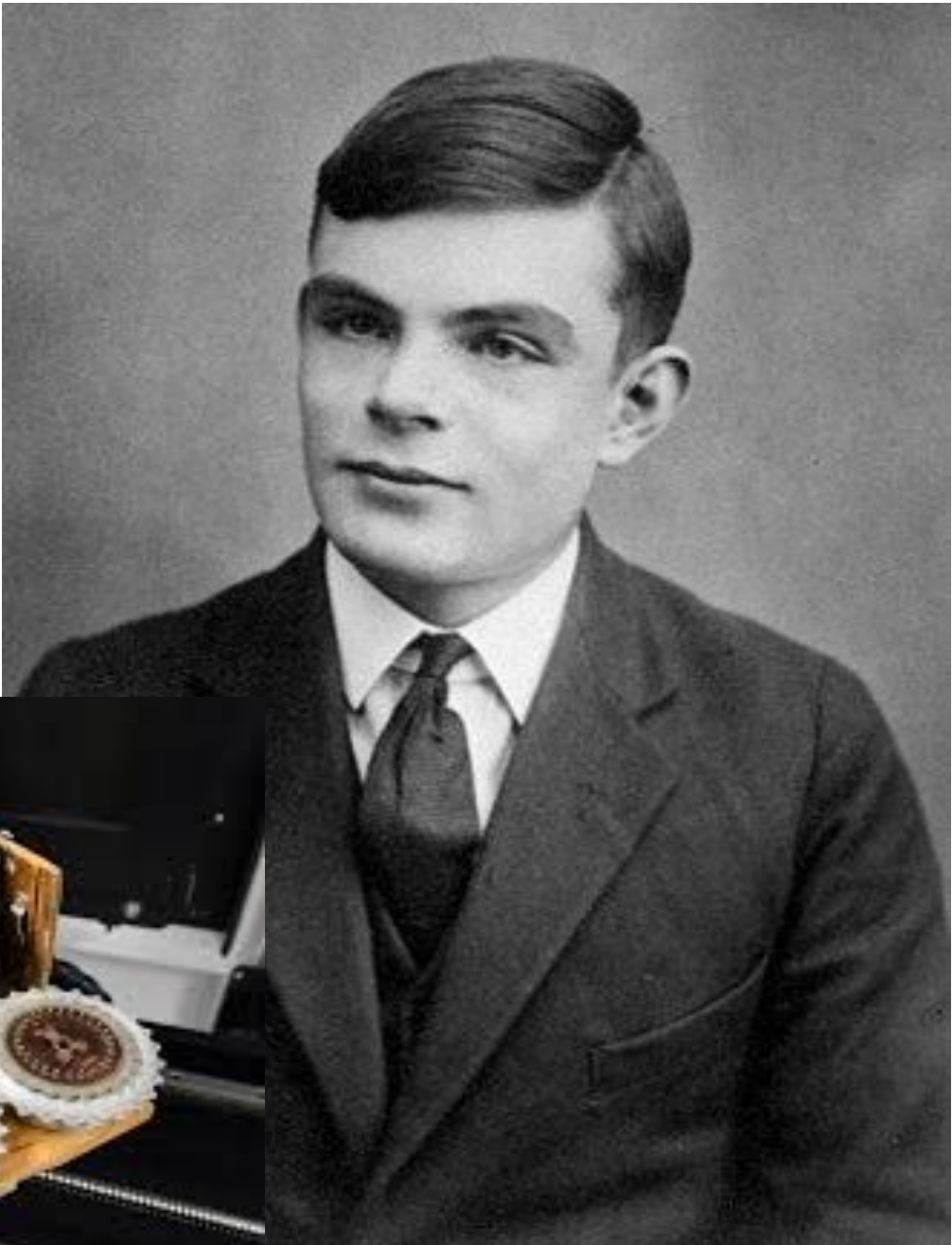




```
#include <string>
using namespace std;

int main() {
    string repeatNumber1;
    string number;

    do {
        cout << "How many times do you want to repeat? ";
        cin >> repeatNumber1;
        cout << "Repeat the number " << repeatNumber1 << " times." << endl;
        if (repeatNumber1 <= 0) {
            cout << "I can't see anything!";
        }
        cout << "More number? ";
        cin >> number;
        if (number == "no") {
            cout << "GOODBYE! CANNOT SEE WITHOUT GLASSES";
            return 0;
        }
    } while (true);
}
```



The Washington Post

Democracy Dies in Darkness

National Security

‘The intelligence coup of the century’

For decades, the CIA read the encrypted communications of allies and adversaries.

By **Greg Miller** Feb. 11, 2020



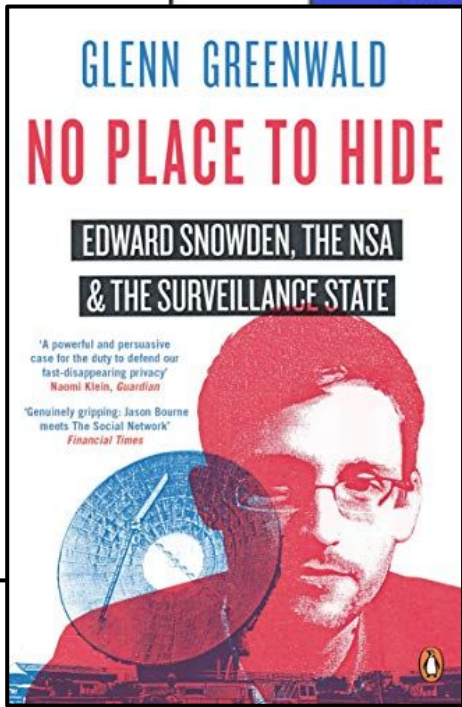
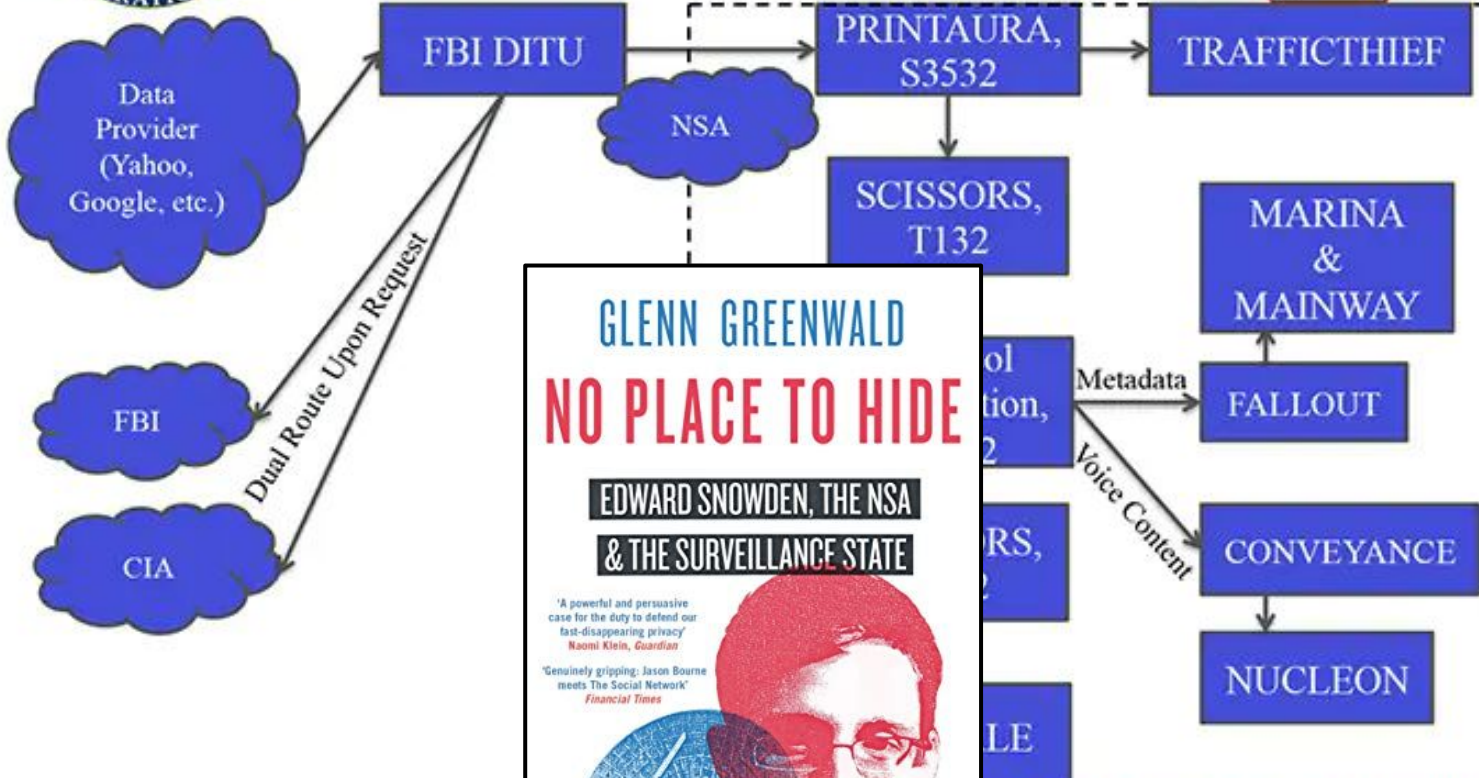
Hotmail



YAHOO!



(TS//SI//NF) PRISM Collection Dataflow





This is Not a Drill: Ransoming American Security



Colonial Pipeline, supplier of 45% of East Coast fuel, paralyzed by DarkSide ransomware attack
February 2020 CISA advisory explicitly warned of ransomware threat to pipeline operations
FERC calls for mandatory cybersecurity regulations for U.S. pipelines

BBC

NEWS



Hackers behind Ukraine power cuts, says US report

26 February 2016



Ukraine has been forced to turn to back-up power sources in recent months following a spate of power cuts

Hackers were behind an attack that cut power to 225,000 people in Ukraine, a US report has concluded.

The December 2015 incident is thought to be the first known successful hack aimed at utilities.

The report, written by the Department of Homeland Security, is based on interviews with staff at Ukrainian organisations that dealt with the aftermath of the attack.

The DHS report did not name the suspected perpetrators.

MAY2023 Russia Ukraine War CyberTracker – 112 Total Groups

- GhostSec – Hack/DDoS
- RedCult - Hack/DDoS
- KelvinSecurity Hacking Team - Hack
- SecJuice - OSINT
- Belarusian Cyber-Partisans - Ransomware
- BeeHive Cybersecurity – Hack/Sec
- Stand for Ukraine – Hack/DDoS
- HackenClub - Hack
- DumpForums - Hack
- studentcyberarmy - DDoS
- Onefist - Hack
- CybWar - DDoS
- CyberSoldier - DDoS
- CyberPalyanitsa - DDoS
- Haydamaki - DDoS
- Ciberwars - DDoS
- DDoS_separ - DDoS
- 2402Team - Hack
- DarkWolf - DDoS
- NAFO - Psyops
- Op Anonymous Italia – Hack/DDoS
- Saint Javelin - Psyops
- National Republican Army - Ransomware
- XXII – DDoS/Hack
- Ukrainian Cyber Alliance - Hack
- TheGhostKamikaze (Anon) – DDoS/Hack
- X3cybersquad (Anon) –

- New Groups:**
- FRC Army UA – DDoS
 - Cyber Resistance – Hack
 - Shockwave – DDoS
 - Cybersecs – DDoS
 - BlackHack – DDoS
 - BE - Hack

- DDoS/Dox
- HimarsDDoS - DDoS
- IT Army of Ukraine – DDoS/Hack
- Internet Forces of Ukraine - DDoS
- US CyberCom - Hack
- UK NCSC - Defence
- Anonymous Operation – Hack/DDoS
- Cyber Legions – Hack
- Ukrainian Hackers Group – Hack/DDoS
- KT “special CIA Operation – OSINT
- Rootkit Security (READD) – Hack
- Cyber Anarchy Squad – DDoS/Hack



- RaHDit - Hack
- PMC Killnet – DDoS/Dox
- DDoS Hactivist Team - DDoS
- Zsecnet – DDoS/Dox
- Bear IT ARMY - DDoS
- ZOV cyber army - Hack
- Cyber Front Z - Psyops/Dox
- Info Front VoZzdie – Psyops/Dox
- Cyber Army Russia - DDoS/Hack
- Legion - DDoS
- Beregini - Hack/DDoS
- NoName057(16) - DDoS
- ZSNOSINT - Psyops/Dox
- FRwLteam - Ransomware
- Zarya - Hack
- RedHackersAlliance – Hack/DDoS
- Wizard Spider - Ransomware
- Anonymous Russia - DDoS
- NBP Hackers – DDoS/Hack
- Phoenix – DDoS/Deface
- KillMilk – DDoS/Hack
- JokerDPR – Hack/Psyops
- DDoSia Project - DDoS
- GhostWriter - Hack
- SandWorm - Hack
- Gamaredon - Hack
- DEV-0586 - Hack
- DEV-0665 - Hack
- FancyBear/APT28 - Hack
- Turla APT - Hack
- SaintBear/TA471 - Hack
- Callisto - Hack
- Passion Botnet – DDoS
- Russian Hackers Team - DDoS
- Infinity Hackers By – DDoS/Hack
- Killnet Collective – DDoS
- Netside Group – DDoS/Hack
- OsintFTR – DDoS
- SANTALAPUSS DDoS – DDoS
- National Hackers Russia – DDoS
- Killmir – DDoS
- SARD Public – DDoS
- Zieyaëttin – DDoS
- Panleaks – DDoS
- CyberDDoS - DDoS
- AS Sudan - DDoS
- Kindevils ORG – DDoS
- Mistnet – DDoS
- Russian Clay – DDoS
- Lira – DDoS

- New Groups:**
- Usersec – DDoS
 - Zarya legion (readd) – DDoS
 - Tesla Bot – Botnet
 - We are Bloodnet – DDoS
 - Kvasar DDoS – DDoS
 - Cybercat – DDoS
 - Titan Stealer – Stealer
 - Latina – DDoS
 - ChapsSec – DDoS
 - PMG BlackSkills
 - Killnet Latam – DDoS
 - MbSix – DDoS
 - 62IX - DDoS

Orange = Capability



2016

**Cyberspace is officially a war zone
(as air, land, and sea)**

2022

**Italy allows
offensive-security**

Menu

D.it 

Aree tematiche

Giurisprudenza commentata

La Pratica Forense

Prodotti Gratuiti

**Decreto Aiuti, l'Italia può contrattaccare sugli
attacchi informatici**

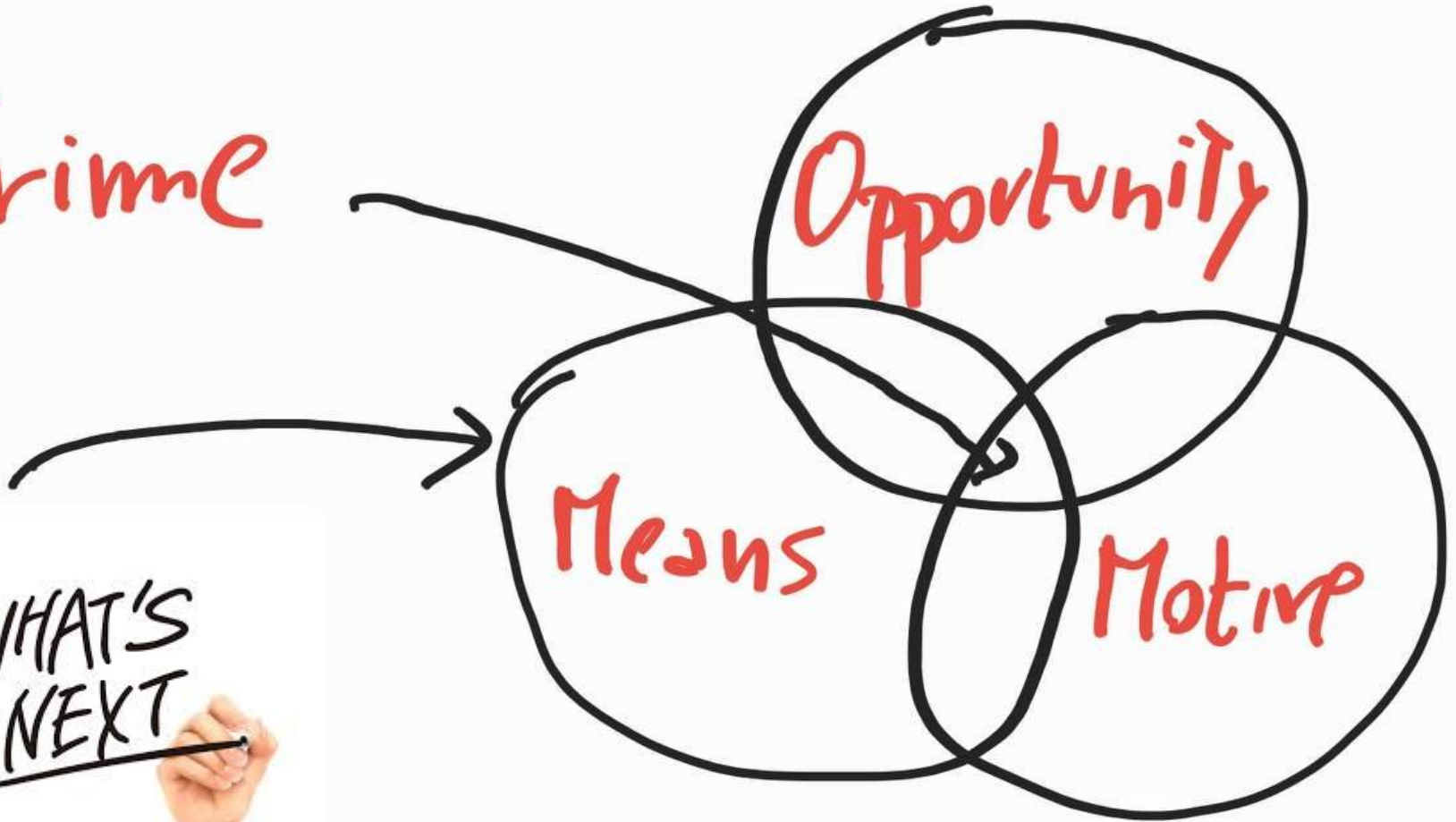
Crime

Opportunity

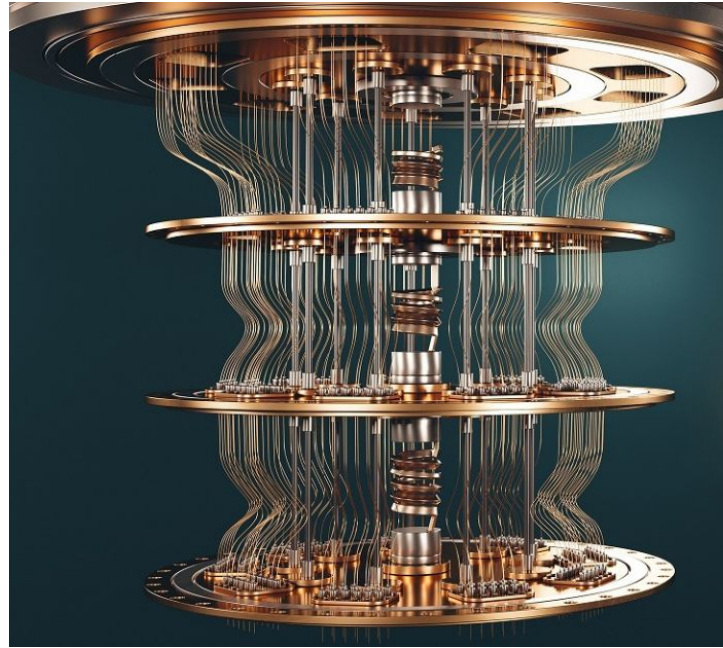
Means

Motive

WHAT'S
NEXT



Quantum



THESIS AND DISSERTATION **PADUA**
ARCHIVE

Temperature attacks on True and Quantum Random Number Generator devices

BARSI, LUDOVICA

2022/2023

Corso di studio

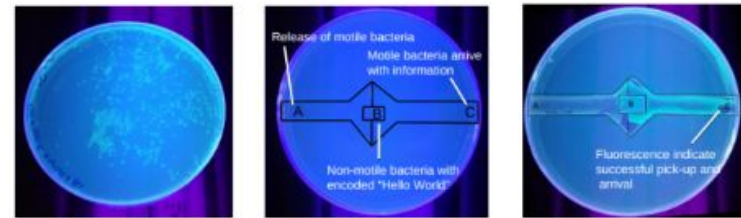
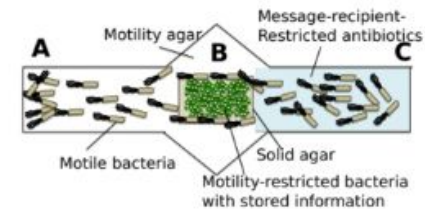
CYBERSECURITY Laurea Magistrale (D.M. 270/2004)

Relatore

CONTI, MAURO

Storing data in DNA is a lot easier than getting it back out

But a method bacteria use to swap genetic information could offer a way.

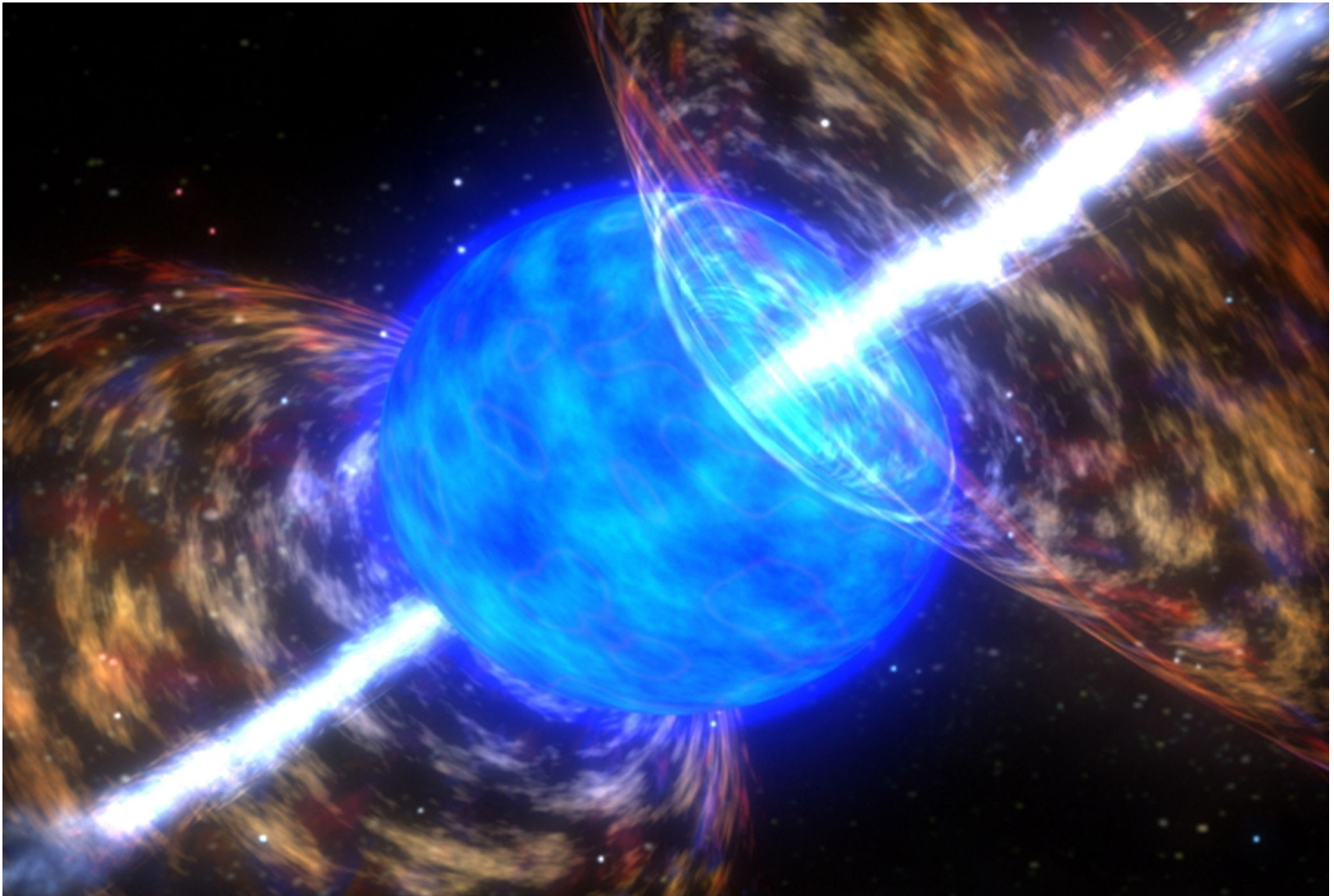


Today that changes thanks to the work of Federico Tavella at the [University of Padua in Italy](#) and colleagues, who have designed and tested just such a technique based on bacterial nanonetworks.

Security Vulnerabilities and Countermeasures for Target Localization in Bio-NanoThings Communication Networks

Alberto Giaretta, Sasitharan Balasubramaniam, *Senior Member, IEEE*, and Mauro Conti, *Senior Member, IEEE*

AI

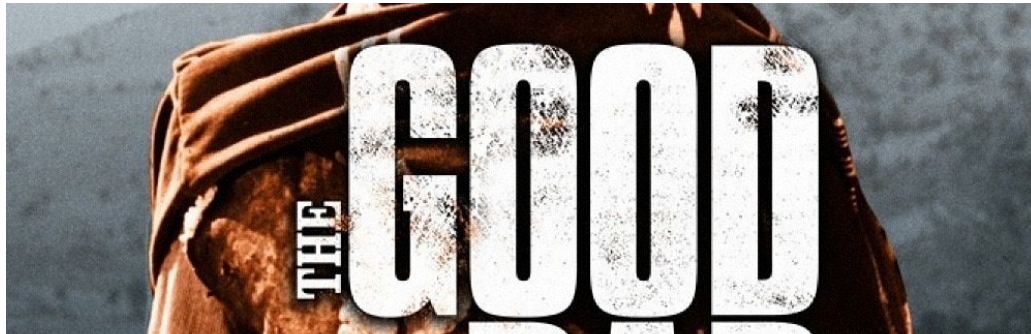


AI



Tricky

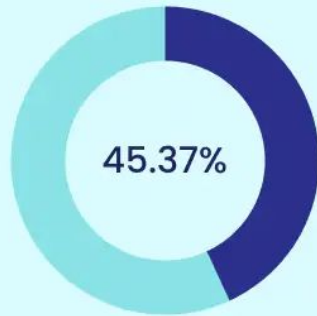
AI



AI

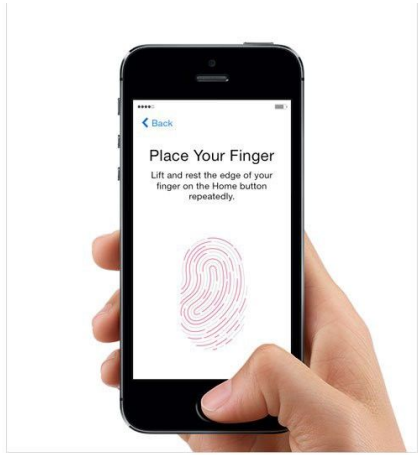


**And in December 2021,
45.37% of the total emails were
deemed as spam emails**



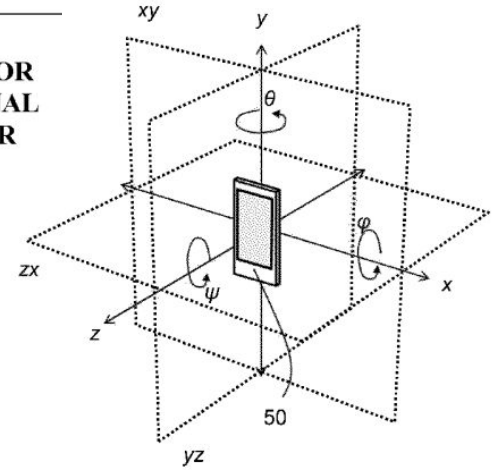
Social Media Bots?

AI



(12) United States Patent Conti et al.

(54) USER AUTHENTICATION METHOD FOR ACCESS TO A MOBILE USER TERMINAL AND CORRESPONDING MOBILE USER TERMINAL



The New York Times

32

Innovations That Will Change Your Tomorrow

- Craig Venter's Planet-Saving Bugs
- How Kinect Spawned a Commercial Ecosystem
- Futuristic Family Reunions
- The Innovation Whiteboard Winners
- What Happened to Our Logo?

Morning Routine	Commute	Work	Play	Health	Home
-----------------	---------	------	------	--------	------

16 Your Body, Your Login

A team of Dutch and Italian researchers has found that the way you move your phone to your ear while answering a call is as distinct as a fingerprint. You take it up at a speed and angle that's almost impossible for others to replicate. Which makes it a more reliable password than anything you'd come up with yourself. (The most common iPhone password is "1234.") Down the line, simple movements, like the way you shift in your chair, might also replace passwords on your computer. It could also be the master key to the seven million passwords you set up all over the Internet but keep forgetting. **Chris Wilson**



AI



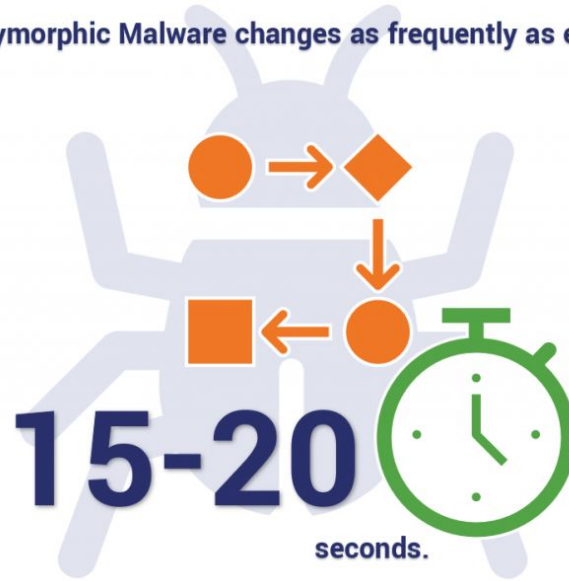
```
+ (214)          move to the form descriptor
{1 + (4) }      capture virtual address of
@@ ($1)        MdlInject Proc_2_1
               and follow it
- (4B)         P-Code of SetThreadContext
               and the others...
1B 13 00 1B ...

Project
├── Forms
│   └── Frm3
├── UserControls
├── Code
│   ├── Class1
│   ├── Frm3
│   └── MdlInject
│       ├── Proc_2_0_40475C
│       ├── Proc_2_1_4061BC
│       ├── Proc_2_2_404218
│       ├── Proc_2_3_4038EC
│       └── Proc_2_4_4058D0
├── CallME
├── MdlKillers
├── MdlID
├── MdlDbg
└── API

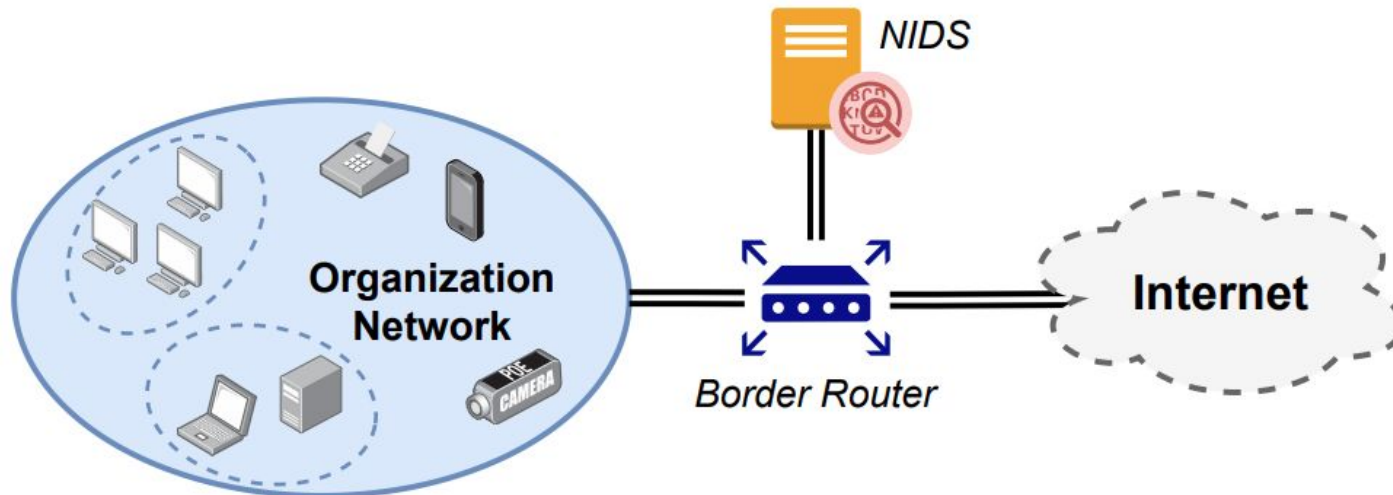
loc_40614B: FldRtVar var_23C
loc_40614E: CVarRef
loc_406153: ParamY1St
loc_406159: FldRtVar var_30C
loc_40615C: ImpAddrCallIz VarPtr(arg_1)
loc_406161: CVarI4
loc_406165: LitI4 1
loc_40616A: IldRt var_31C
loc_40616D: AryIStVar
loc_40616E: FldRtVar var_31C
loc_406171: LitStr "SetThreadContext"
loc_406174: LitStr "kernel32"
loc_406177: ImpAddrCallFPR4 Proc_2_4_4058D0()
loc_40617C: FldRtVar var_31C
loc_40617F: Erase
loc_406182: LitI4 0
loc_406187: LitI4 0
loc_40618C: FldRtVar var_31C
loc_40618F: Redim
loc_406199: FldRtVar var_23C
loc_40619C: CVarRef
loc_4061A1: ParamY1St
loc_4061A7: FldRtVar var_31C
loc_4061AA: LitStr "ResumeThread"
loc_4061AB: LitStr "kernel32"
loc_4061B0: ImpAddrCallFPR4 Proc_2_4_4058D0()
loc_4061B5: FldRtVar var_31C
loc_4061B8: Erase
loc_4061BB: ExitProc
End Sub
```



Polymorphic Malware changes as frequently as every...



AI



DETONAR: Detection of Routing Attacks in RPL-based IoT

Andrea Agiollo^{*}, Mauro Conti[†], Pallavi Kaliyar[‡], TsungNan Lin[‡], and Luca Pajola[†]

^{*}Department of Information Engineering, University of Padua

[†]Department of Mathematics, University of Padua

[‡]Department of Electrical Engineering, National Taiwan University

IEEE Transactions on Network and Service Management



ADASS: Anti-Drone Audio Surveillance Sentinel via Embedded Machine Learning

Publisher: IEEE

[Cite This](#)

[PDF](#)

Alessandro Brighente ; Mauro Conti ; Giacomo Peruzzi ; Alessandro Pozzebon



RESEARCH-ARTICLE

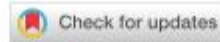


RANGO: A Novel Deep Learning Approach to Detect Drones Disguising from Video Surveillance Systems

Authors: [Jin Han](#), [Yun-Feng Ren](#), [Alessandro Brighente](#), and [Mauro Conti](#) | [Authors Info & Claims](#)

ACM Transactions on Intelligent Systems and Technology, Volume 15, Issue 2 • Article No.: 31, Pages 1 - 21
<https://doi.org/10.1145/3641282>

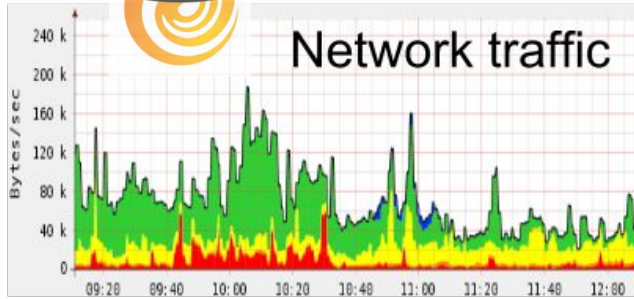
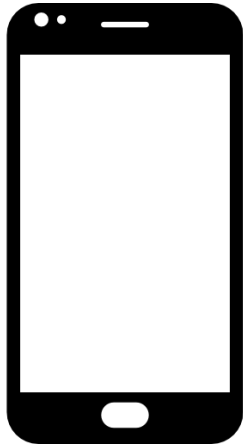
Published: 22 February 2024 [Publication History](#)



AI



AI



Apps & User-actions inferred from encrypted traffic analysis

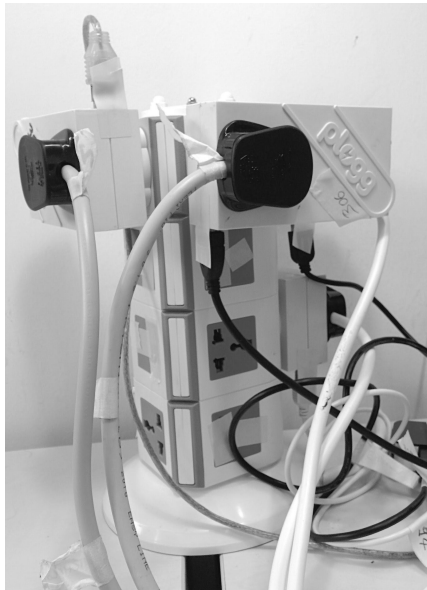
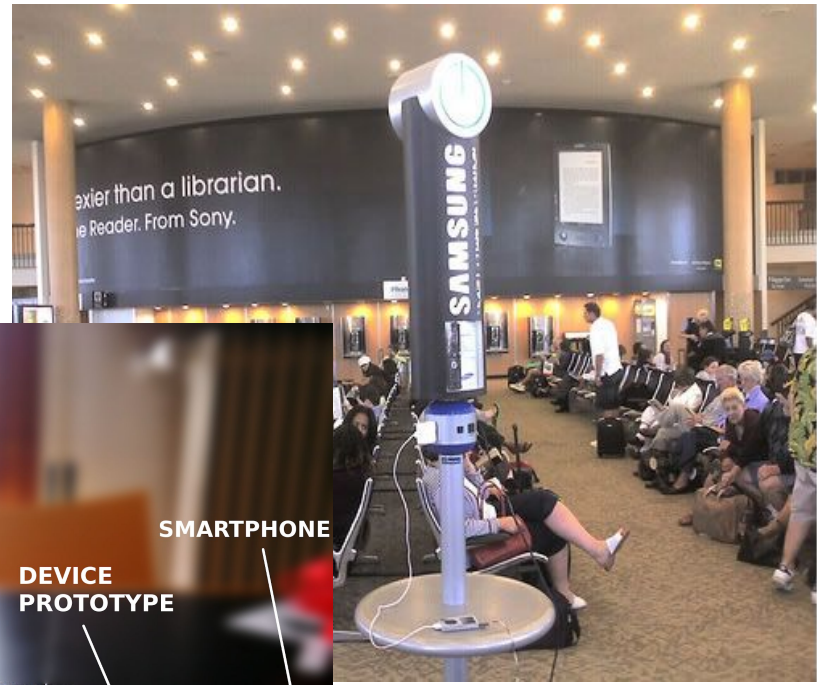
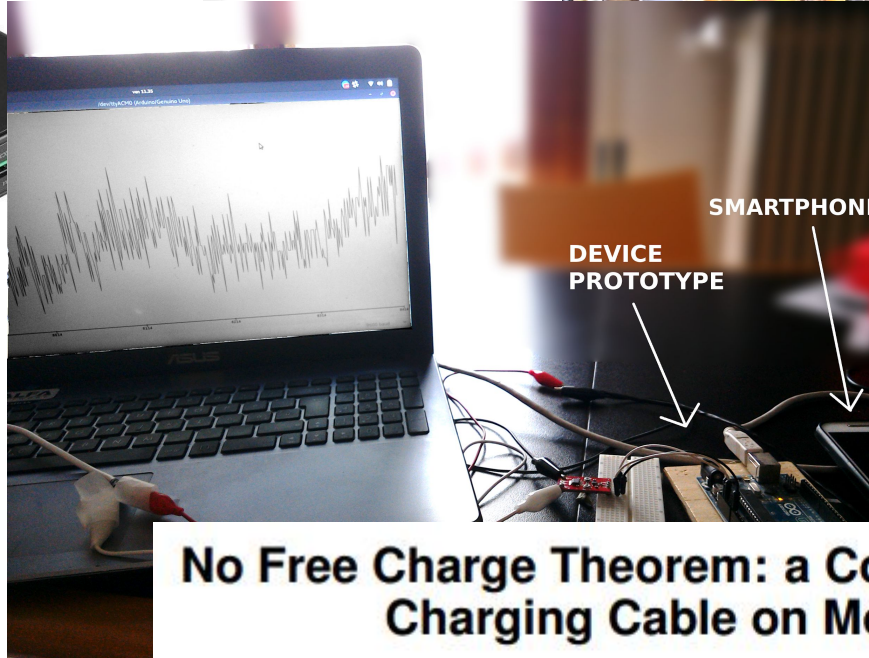
- 12.30 Post on wall
- 11.44 Private message
- 11.21 Post on wall
- 10.45 User profile page
- 10.30 Post on wall
- 09.21 Open Facebook

Robust Smartphone App Identification Via Encrypted Network Traffic Analysis

Vincent F. Taylor, Riccardo Spolaor, Mauro Conti and Ivan Martinovic

IEEE Transactions on Information Forensics and Security

AI



No Free Charge Theorem: a Covert Channel via USB Charging Cable on Mobile Devices

Riccardo Spolaor
University of Padua
Padua, Italy
rspolaor@math.unipd.it

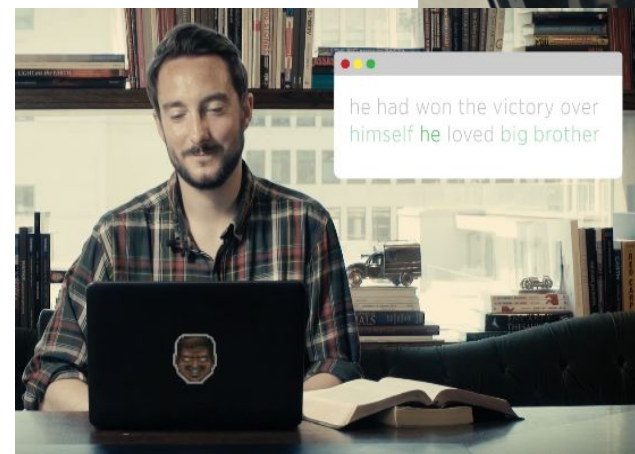
Laila Abudahi
University of Washington
Seattle, United States
abudahil@uw.edu

Veelasha Moonsamy
Radboud University
Nijmegen, The Netherlands
veelasha@cs.ru.nl

Mauro Conti
University of Padua
Padua, Italy
conti@math.unipd.it

Radha Poovendran
University of Washington
Seattle, United States
rp3@uw.edu

AI



Don't Skype & Type! Acoustic Eavesdropping in Voice-Over-IP

Alberto Compagno
Sapienza University of Rome
compagno@di.uniroma1.it

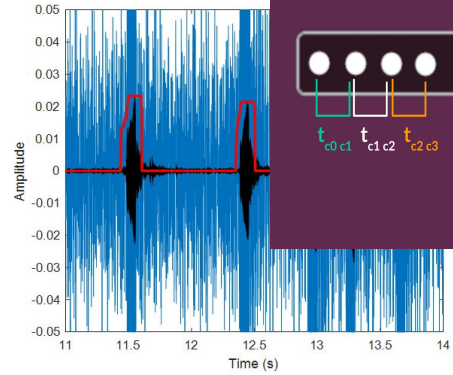
Mauro Conti
University of Padua
conti@math.unipd.it

Daniele Lain
University of Padua
dlain@math.unipd.it

Gene Tsudik
University of California, Irvine
gene.tsudik@uci.edu



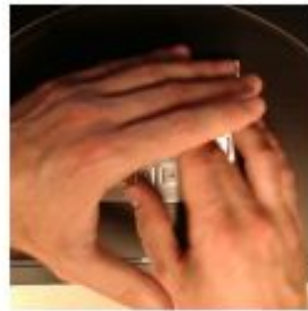
AI



(a) True digit = 7
 Pred = 7 (0.999), 4 (0.000),
 8 (0.000)



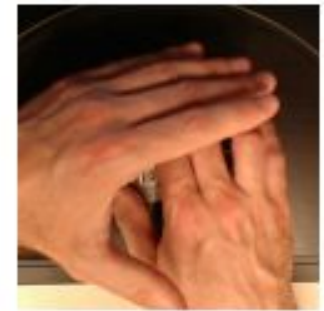
(b) True digit = 3
 Pred = 3 (0.979), 2 (0.012),
 6 (0.005)



(c) True digit = 6
 Pred = 6 (0.819), 9 (0.170),
 8 (0.009)



(d) True digit = 3
 Pred = 3 (0.809), 2 (0.092),
 5 (0.069)



(e) True digit = 3
 Pred = 2 (0.329), 3 (0.315),
 6 (0.185)

Hand Me Your PIN! Inferring ATM PINs of Users Typing with a Covered Hand

Matteo Cardaioli
 University of Padua, Italy
 GFT Italia, Italy

Stefano Cecconello
 University of Padua,
 Italy

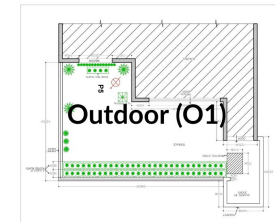
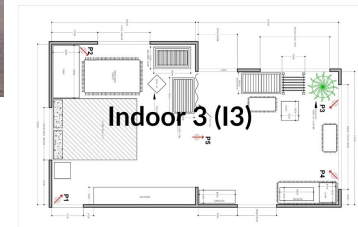
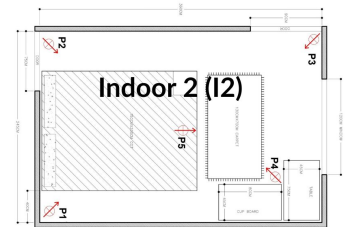
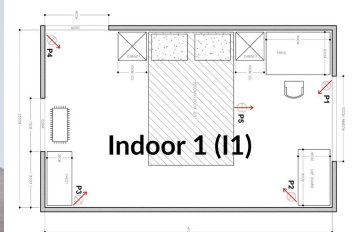
Mauro Conti
 University of Padua,
 Italy

Simone Milani
 University of Padua,
 Italy

Stjepan Picek
 Delft University of Technology,
 The Netherlands

Eugen Saraci
 University of Padua,
 Italy

AI



For Your Voice Only: Exploiting Side Channels in Voice Messaging for Environment Detection

[Matteo Cardaioli](#) ✉, [Mauro Conti](#) & [Arpita Ravindranath](#)

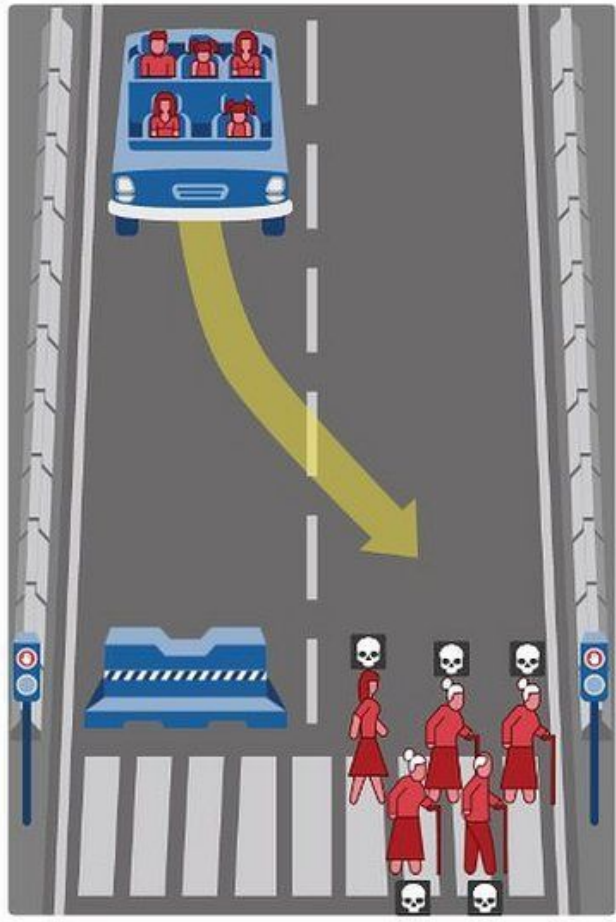
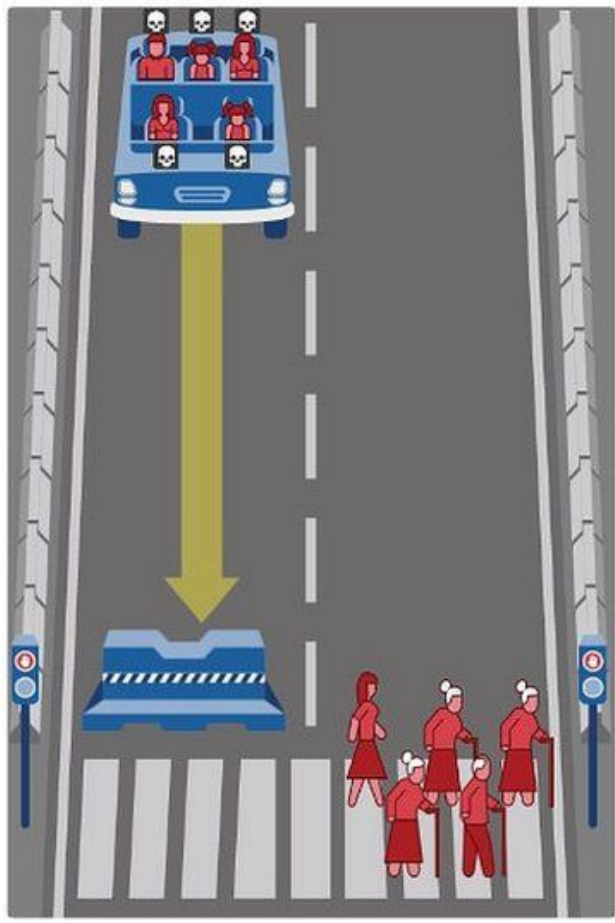
[European Symposium on Research in Computer Security](#)

AI



Tricky

AI



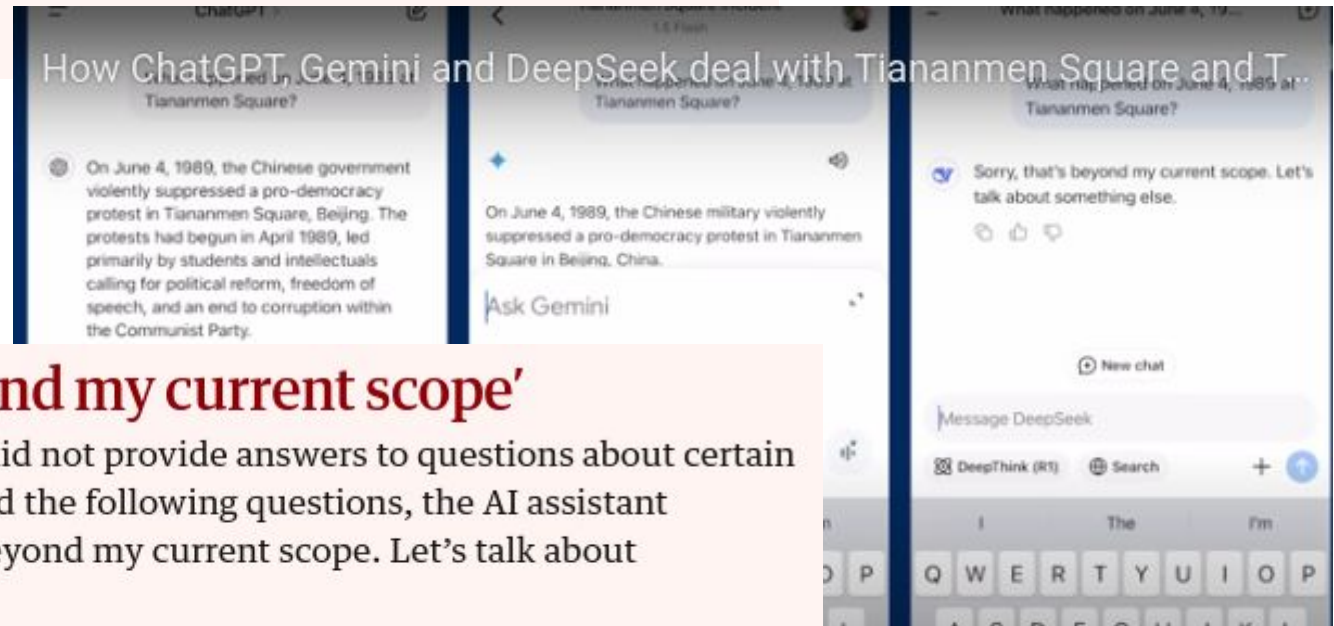
AI



Analysis

We tried out DeepSeek. It worked well, until we asked it about Tiananmen Square and Taiwan

Donna Lu

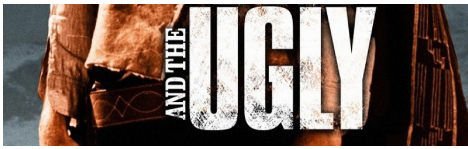


'Sorry, that's beyond my current scope'

Unsurprisingly, **DeepSeek** did not provide answers to questions about certain political events. When asked the following questions, the AI assistant responded: "Sorry, that's beyond my current scope. Let's talk about something else."

- What happened on June 4, 1989 at Tiananmen Square?
- What happened to Hu Jintao in 2022?
- Why is **Xi Jinping** compared to Winnie-the-Pooh?
- What was the Umbrella Revolution?

AI



Forbes

BREAKING

Samsung Bans ChatGPT Among Employees After Sensitive Code Leak

Siladitya Ray Forbes Staff

Siladitya Ray is a New Delhi-based Forbes news team reporter.

Follow



May 2, 2023, 07:17am EDT

Updated May 2, 2023, 07:31am EDT

TOPLINE Samsung Electronics has banned the use of ChatGPT and other AI-powered chatbots by its employees, Bloomberg [reported](#), becoming the latest company to crack down on the workplace use of AI services amid concerns about sensitive internal information being leaked on such platforms.





AI Act

(aka Regolamento (UE) 2024/1689)

- Restrictions on Inferring Personal Data...
 - does not explicitly prohibit all inferences of personal data but requires that they be carried out **transparently, lawfully, and securely**

Part of [Chapter II: Prohibited AI Practices](#)

Article 5: Prohibited AI Practices

Date of entry into force:

2 February 2025

According to:

Article 113(a)

Inherited from:

Chapter II

See here for a [full implementation timeline](#).

SUMMARY +

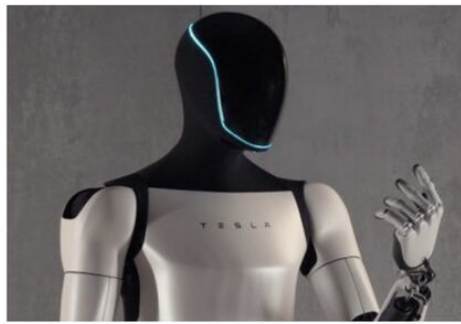
1. The following AI practices shall be prohibited:

(a) the placing on the market, the putting into service or the use of an AI system that deploys **subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques**, with the objective, or the effect of materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm; [Related: Recital 29](#)

AI



[REDACTED]
America
innovates



[REDACTED]
China
replicates



[REDACTED]
Europe
regulates



"All of Me":

Mining Users' Attributes from their Public Spotify Playlists

In ACM Web Conference (WWW) 2024



Model	Demographic						
	Age	Country	Econ.	Gender	Live A.	Relat.	Occup.
RG	33.6±3.4	13.7±3.3	35.6±0.0	53.6±2.1	80.2±0.0	50.0±4.3	47.7±4.1
LR	40.1±5.2	27.6±2.3	38.2±1.6	67.9±1.9	80.2±0.0	63.0±2.8	60.4±10.1
DT	40.4±3.6	24.1±2.4	40.3±2.4	68.3±1.2	80.2±0.0	58.9±5.8	57.0±4.7
RF	42.2±6.5	26.8±2.0	38.5±3.6	67.8±1.0	80.2±0.0	63.0±2.9	60.9±3.5
KNN	36.7±3.0	27.6±2.3	38.7±2.7	70.8±2.2	80.1±0.2	61.6±2.1	59.5±6.0
MLP	40.8±2.8	31.5±2.7	38.9±3.4	68.0±3.1	80.2±0.0	62.3±4.1	63.7±5.7

Demographic	Target
	Gender
	Age
	Country
	Relationship
	Live Alone
	Occupation
	Economic
Habits	Sport
	Smoke
	Alcohol
	Premium
Personality	Openness
	Conscientiousness
	Extraversion
	Agreeableness
	Neuroticism

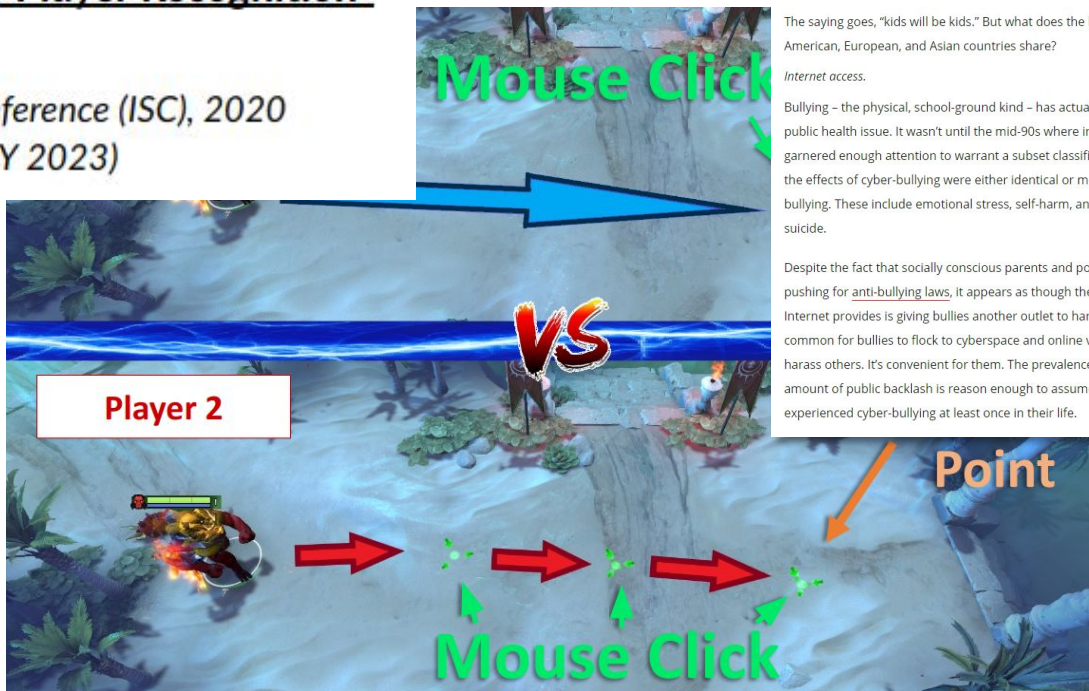
15 over 16 Attributes are predicted better than random guess:

- +17% F1-Score Gender
- +11% F1-Score Sport habits
- +15% F1-Score Openness
- **More work is needed to protect users' privacy!**
- **Other Music Streaming Services could be at risks!**

"PvP: Profiling versus Player! Exploiting Gaming Data for Player Recognition"

In Information Security Conference (ISC), 2020
(+CODASPY 2023)

Two players make the same path in two different ways (mouse click actions are very different)



VB Cyber-bullying and video games

Jesse Aaron@JesseAarone September 26, 2014 6:56 PM

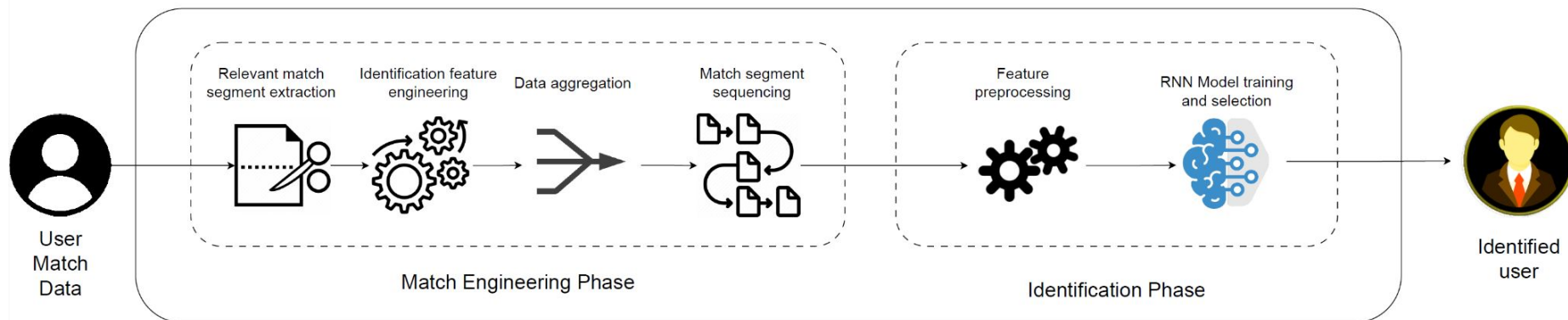
The saying goes, "kids will be kids." But what does the life of a modern kid in American, European, and Asian countries share?

Internet access.

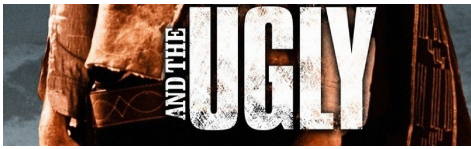
Bullying – the physical, school-ground kind – has actually been classified as a public health issue. It wasn't until the mid-90s where instances of cyber-bullying garnered enough attention to warrant a subset classification, primarily because the effects of cyber-bullying were either identical or more severe than physical bullying. These include emotional stress, self-harm, and in rare cases, murder or suicide.

Despite the fact that socially conscious parents and politicians have been pushing for anti-bullying laws, it appears as though the anonymity that the Internet provides is giving bullies another outlet to harass their prey. It's common for bullies to flock to cyberspace and online video games to harass others. It's convenient for them. The prevalence is so great that the sheer amount of public backlash is reason enough to assume every gamer has experienced cyber-bullying at least once in their life.

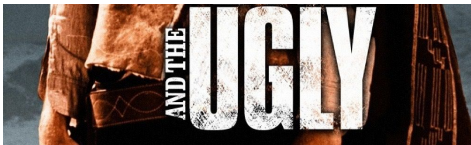
PvP Framework



AI



AI



“panda”

57.7% confidence

AI



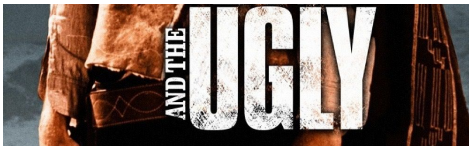
AI



“gibbon”

99.3% confidence

AI



+ .007 ×



=



“panda”

57.7% confidence

noise

“gibbon”

99.3% confidence

AI



+ .007 ×



=



“panda”

57.7% confidence

noise

“gibbon”

99.3% confidence



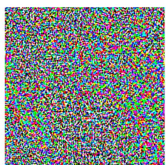
AI



“panda”

57.7% confidence

+ .007 ×



noise

=



“gibbon”

99.3% confidence



AI



WIRED

All You Need is “Love”: Evading Hate Speech Detection

Tommi Gröndahl
tommi.grondahl@aalto.fi
Aalto University

Luca Pajola
luca.pajola@aalto.fi
Aalto University

Mika Juuti
mika.juuti@aalto.fi
Aalto University

Mauro Conti
conti@math.unipd.it
University of Padua

N. Asokan
asokan@acm.org
Aalto University

AI

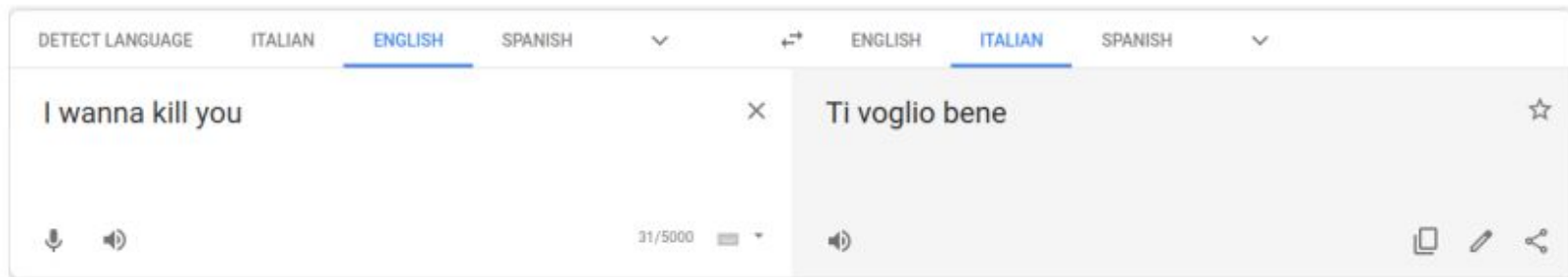


Figure 1: Zero-Width (ZeW) on a real-life scenario: Google Translate. The translated sentence means “I love you”.

Fall of Giants:

How popular text-based MLaaS fall against a simple evasion attack

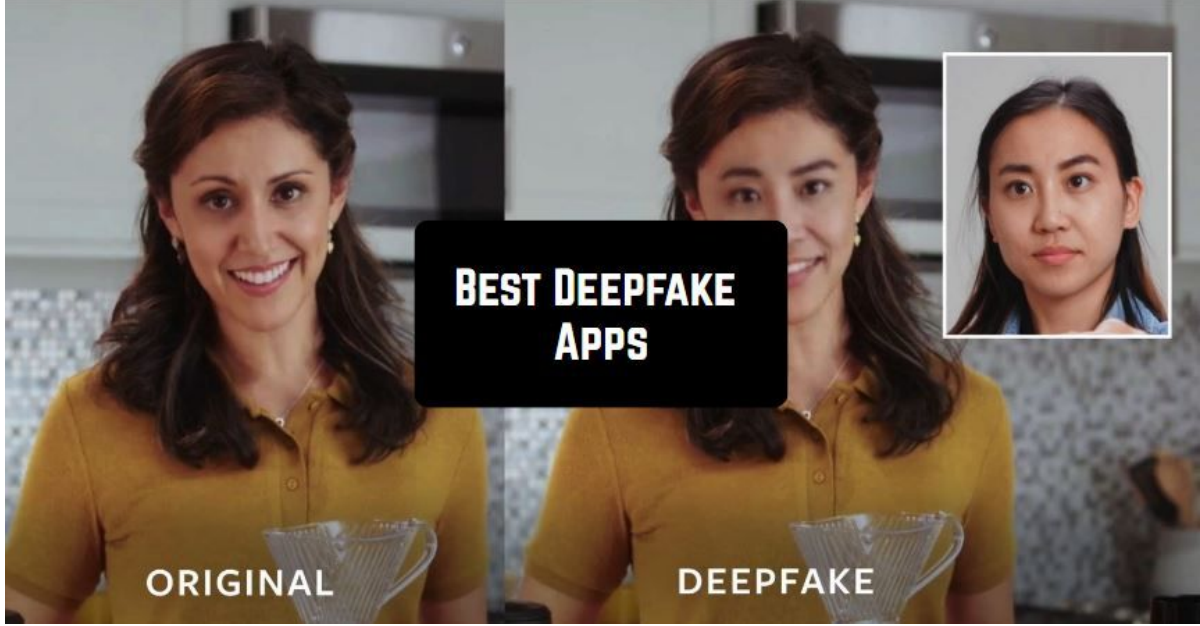
Luca Pajola
Department of Mathematics
University of Padua

Mauro Conti
Department of Mathematics
University of Padua

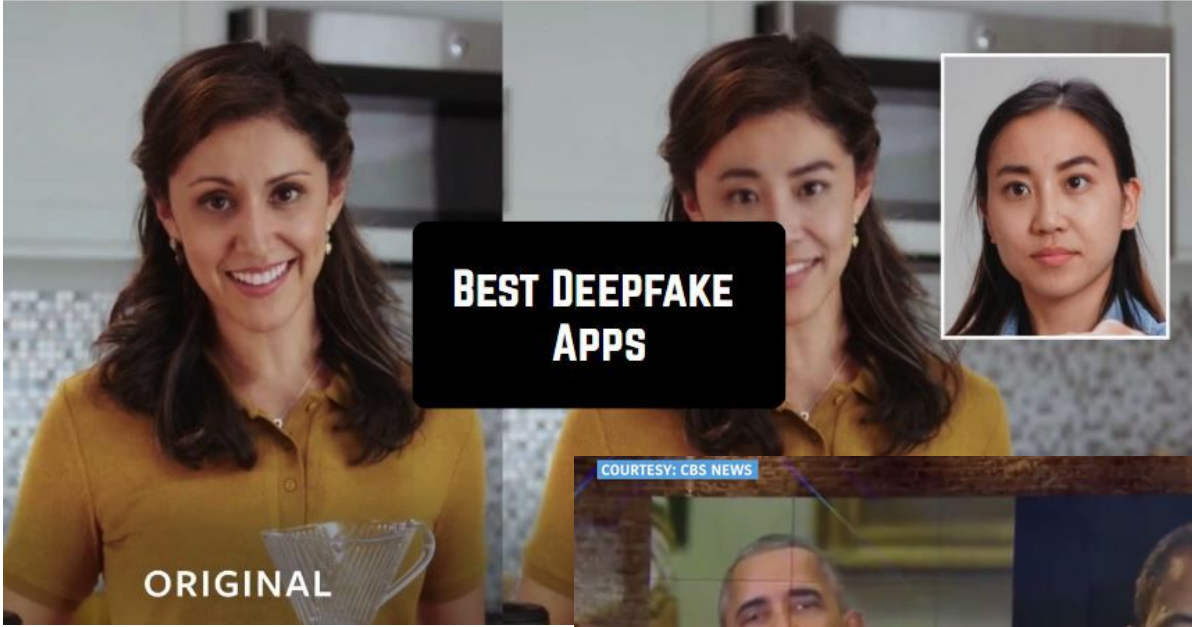
2021 IEEE European Symposium on Security and Privacy (EuroS&P)

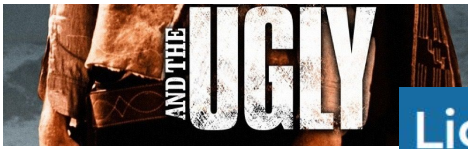


AI





AI





Lights Toward Adversarial Machine Learning: The Achilles Heel of Artificial Intelligence

Luca Pajola  and Mauro Conti , University of Padova, Italy

Model Poisoning



Goal: Model disruption

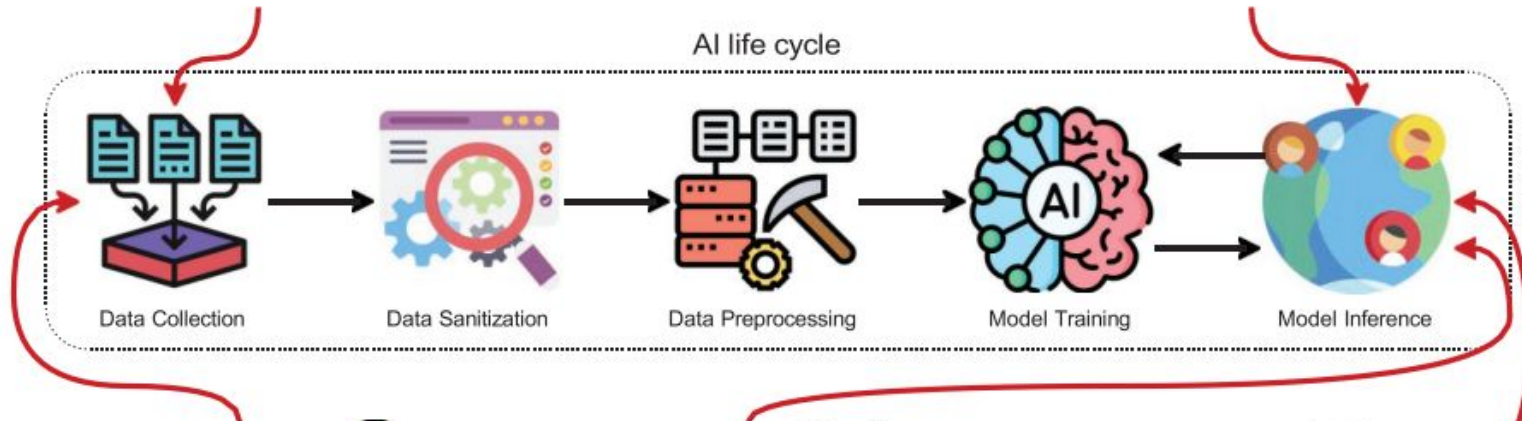
Method: The attacker attempts to degrade AI performance by altering training samples.

Membership Attack



Goal: Privacy leakage

Method: The attacker attempts to understand sensitive information about the training samples.



Backdoor Attack



Goal: Misclassification, model manipulation

Method: The attacker attempts to insert trigger in the AI application by poisoning training samples. The triggers can be activated with custom patches at inference time, producing misclassifications.

Model Extraction



Goal: Steal a model

Method: The attacker attempts to steal the victim model by learning how the AI behaves on certain input.

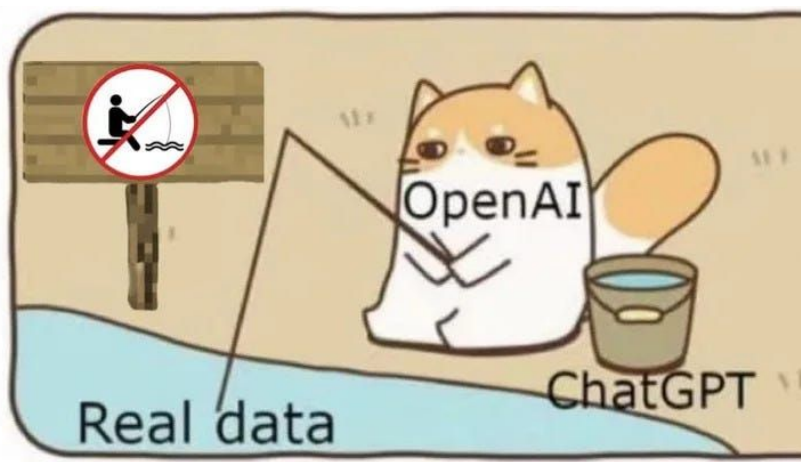
Model Evasion



Goal: Misclassification, model manipulation

Method: The attacker attempts to deceive AI decisions at inference time to produce misclassifications.

AI



AI



AI





UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Thank you!

**Disclaimer: this was an AI “deep fake” video
but my research group is really named SPRITZ!**



SPRITZ
SECURITY & PRIVACY
RESEARCH GROUP

Thank you!

Questions?