Co-funded by the
Erasmus+ Programme
of the European Union

Pathway in Enterprise Systems Engineering (PENS)

Trust, Artificial Intelligence and Cybersecurity

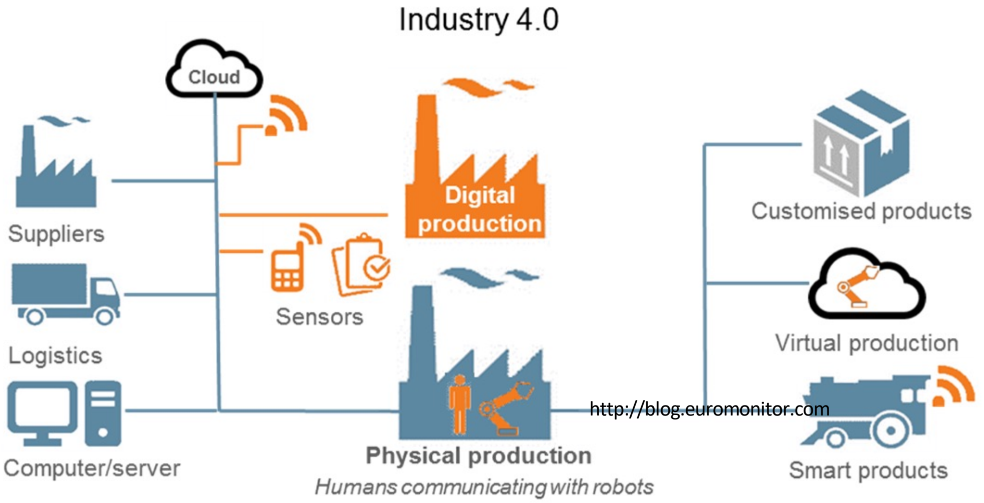
Giorgio Giacinto

February, 21st 2019

London

Hyperconnected world

- Today any *object* can be connected to any other object
  - Data acquisition, data sharing
  - Remote command and control
- Enterprise Systems
  - Traditional Administrative Tasks
  - Customer Relationship
  - Production / Sensors
  - Supply chain
  - *Cloud* storage



- The web of connections is a complex systems
  - The defence of the system requires securing the entire attack surface
  - …but an adversary only needs finding just one vulnerability

Digital Transformation

- Increasing portions of our **daily lives** are managed by software artefacts
- Increasing portions of **enterprise tasks** are managed by software artefacts
  - Data exchange and sharing
  - Cyber-physical systems
    - Kinetic activities depend on the results of data processing
    - Actions depending on (big) data from multiple sources (IoT)
  - Interactions between different software modules
  - Interactions between humans and machines through software
  - Systems accessed from multiple entry points, networks…

Co-funded by the Erasmus+ Programme of the European Union

It is «about trust»

Robert De Niro and Ben Stiller in "Meet the Parents"

Co-funded by the Erasmus+ Programme of the European Union

Trust in the cyber «virtual» world

- In the cyberspace, trust relationships can be established
  - among persons
  - among devices
  - among software modules

  with well defined trust boundaries

- Danger: The cyberspace makes it easy to trust someone or something even with few evidences
- Beware: Trust relationships cannot be considered as transitive
  - an entity that is a member in different relationships, does not cause other entities in the pairs to share trust
  - very difficult, and often impossible to completely check

Enteprise software, people and trust

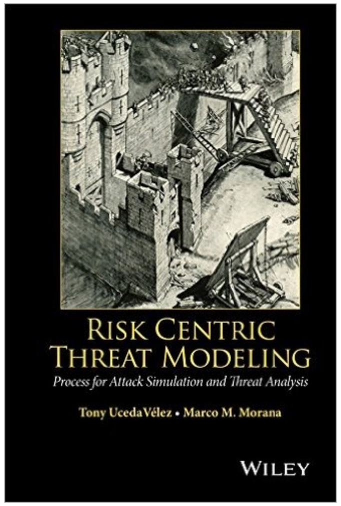
- Enterprise software is a system of modules
  - they interact to ask for / provide services
  - they should trust each other
- The protocol for data exchange, service requests etc. should be designed in order to assure that the appropriate level of trust is verified
- Employees' desktop computers must be carefully configured to ensure that trust is always enforced
  - the desktop machine is clean?
  - are we using adequate authentication mechanisms?
  - etc.

Entering the circle of trust

Threat modelling

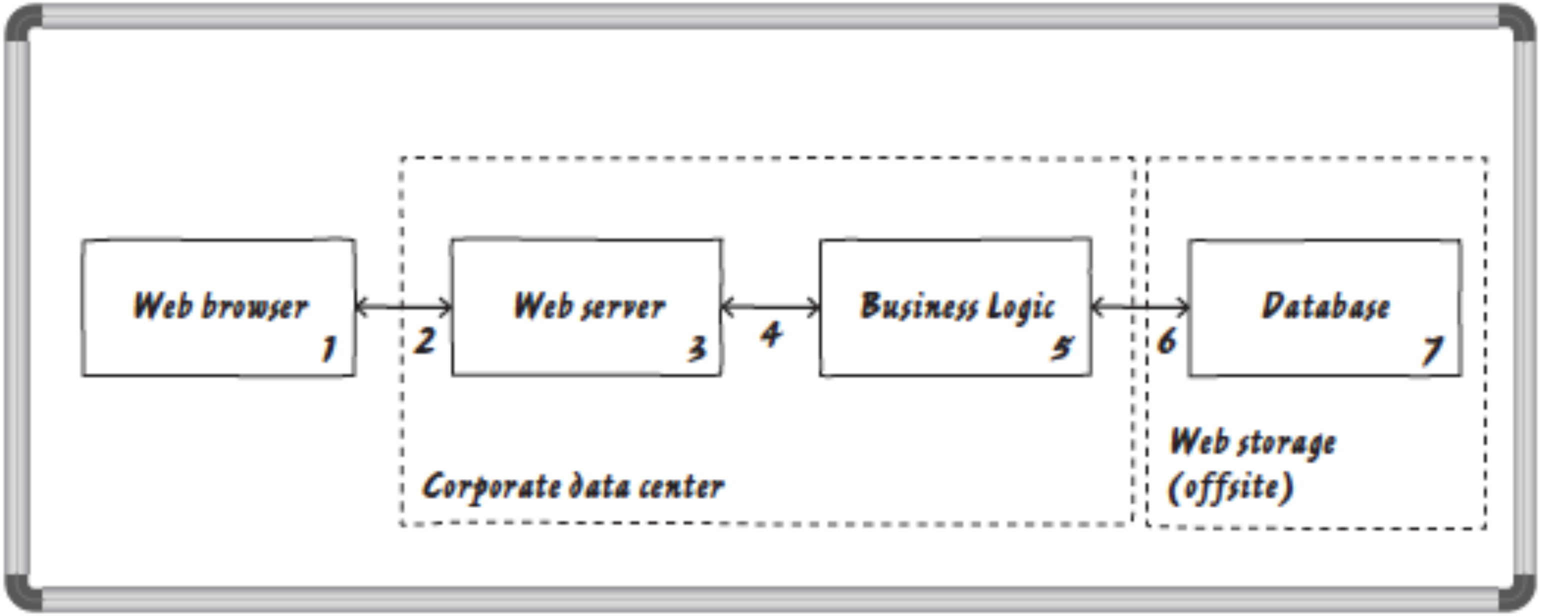
[Application] Threat Modeling – a strategic process aimed at considering possible attack scenarios and vulnerabilities within a proposed or existing application environment for the purpose of clearly identifying risk and impact levels

Tony Uceda Velez and Marco M. Morana, *Risk Centric Threat Modeling*, 2015



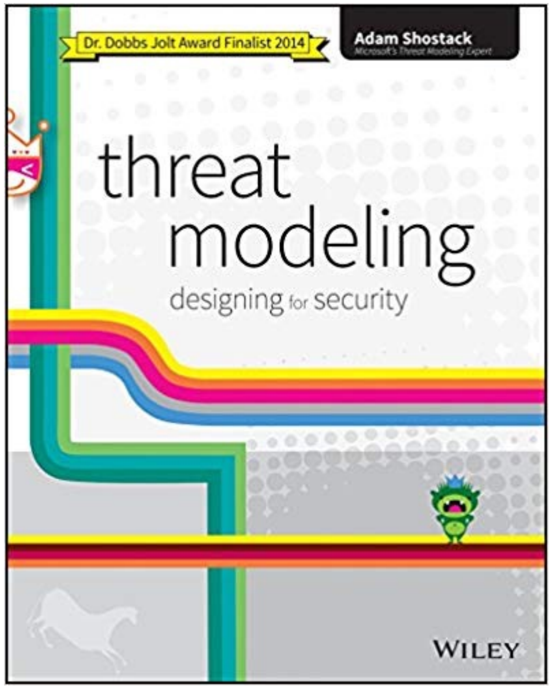
Model the system

- Graphical sketches
- Identification of *Trust Boundaries*



What can go wrong?

- **STRIDE** *taxonomy*
  - Spoofing
  - Tampering
  - Repudiation
  - Information Disclosure
  - Denial of Service
  - Elevation of Privilege



STRIDE

THREAT	PROPERTY VIOLATED	TYPICAL VICTIM
Spoofing	Authentication	Processes External entities People
Tampering	Integrity	Processes Data stores Data flows
Repudiation	Non-Repudiation	Processes
Information Disclosure	Confidentiality	Processes Data stores Data flows
Denial of Service	Availability	Processes Data stores Data flows
Elevation of Privilege	Authorization	Processes

Trustworthiness

"Cute and Cuddly"

Attack kill-chain

https://cloudblogs.microsoft.com/microsoftsecure/2018/12/03/
analysis-of-cyberattack-on-u-s-think-tanks-non-profits-public-sector-by-unidentified-attackers/

Artificial Intelligence and Cybersecurity

Attack

Track and model the behaviour of the *victim* in order to craft *targeted* social engineering attacks.

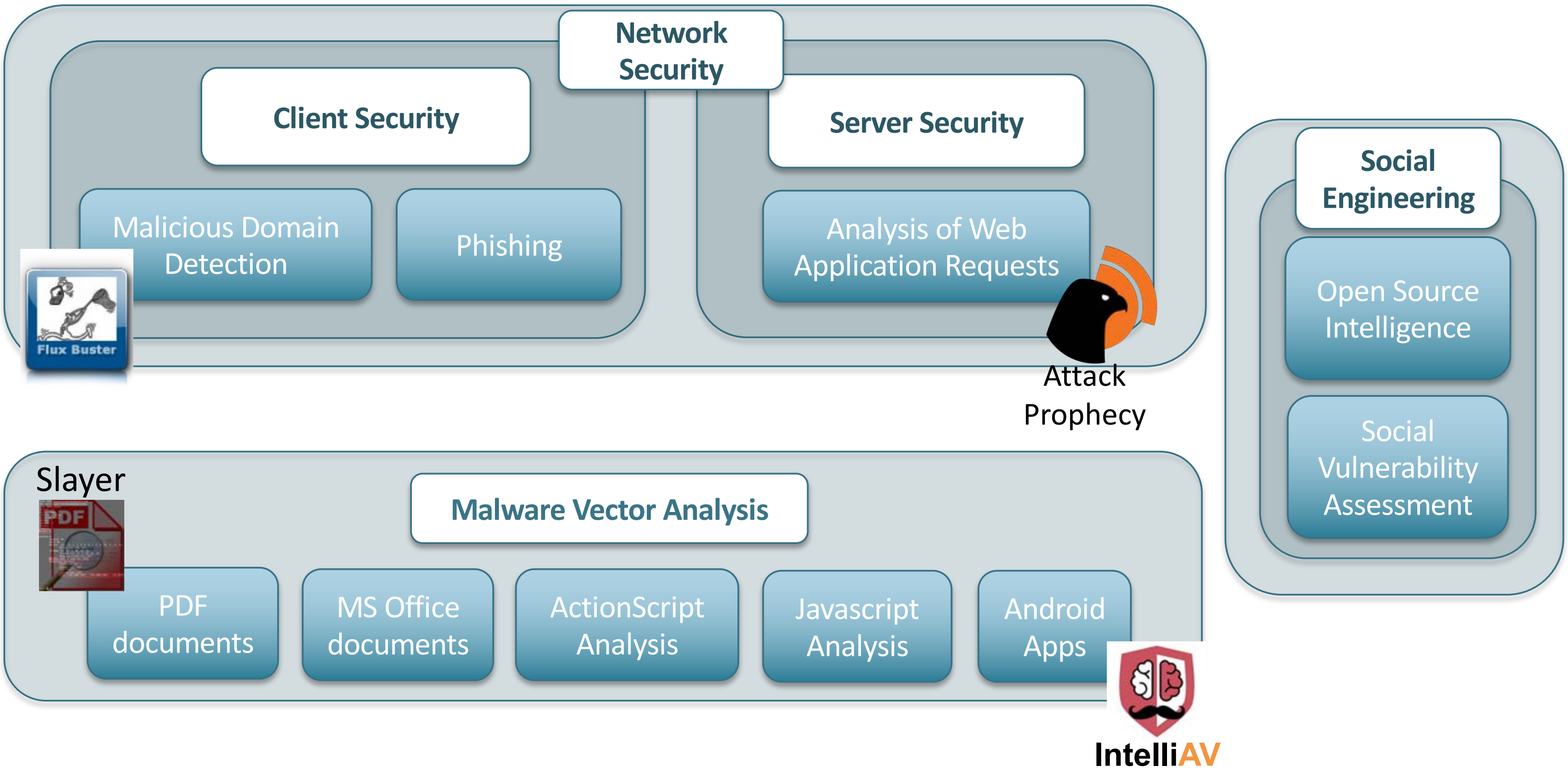
Discover vulnerabilities in networks' and systems' configurations, and in any software module in the target system.
Creation of polymorphic malware samples.

Defence

Track and model the behaviour of attackers to design and implement effective strategies for protection and defence (e.g., firewalls, blacklists, etc.).

Analysis of web applications, software modules, and documents for the early detection of vulnerabilities or malicious components.

Examples of activities to break the kill chain



Can Artificial Intelligence Be Secure?

Artificial Intelligence

The availability of large amounts of **data** from **multiple**, **interconnected** objects and sensors is the driver for a wide adoption of AI

Recommendation Systems and Deep Learning @ eBay

AI & Cybersecurity

- Capability of dealing with vast amount of data
  - **Complexity** of AI algorithms (e.g., deep learning)
  - **Trust** in the implementation

- Interpretability of AI algorithms
  - Complexity increases the likelihood of **vulnerabilities**
  - **Safety & Security** require transparency

- Interconnection
  - Possibility for **Maliciously Targeting AI algorithms** from a **remote** location to **disrupt logical or physical systems**

Need for
human supervision
from design to decision

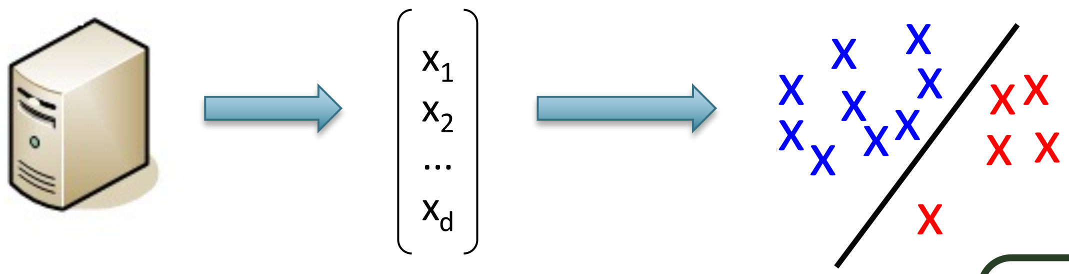
Co-funded by the Erasmus+ Programme of the European Union

Learning & Intelligence

- Building machines that can automatically perform *tedious* classification tasks *with high accuracy*.

TRAINING DATA → FEATURE REPRESENTATION → LEARNING ALGORITHM e.g., SVM, DNN

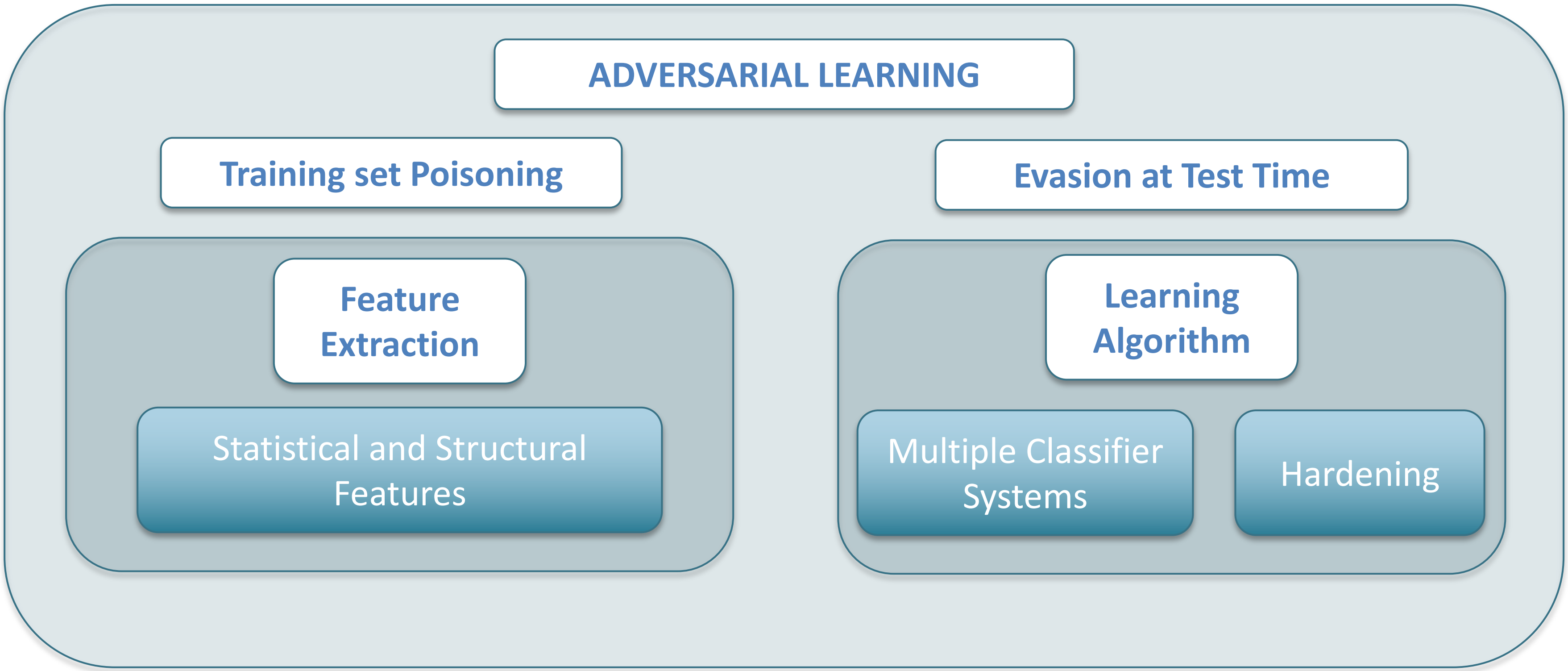
- Learning algorithm
- Parameters (e.g. feature weights)
- Feedback on decisions



Prior definition of **data classes**

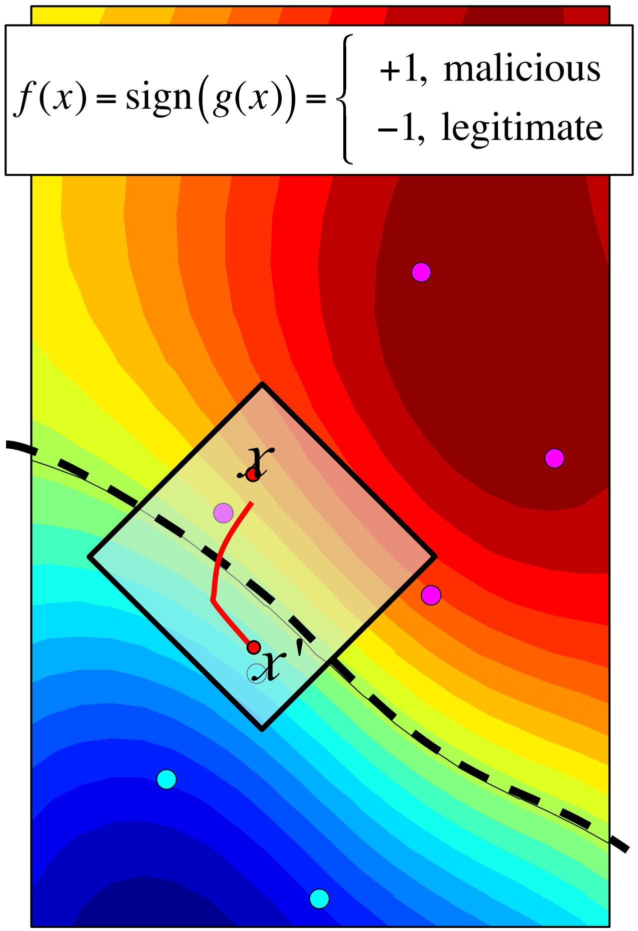
The **learning algorithm** is based on **the goal of learning** - **the objective function** -

Adversarial (Machine) Learning



Evasion Attacks

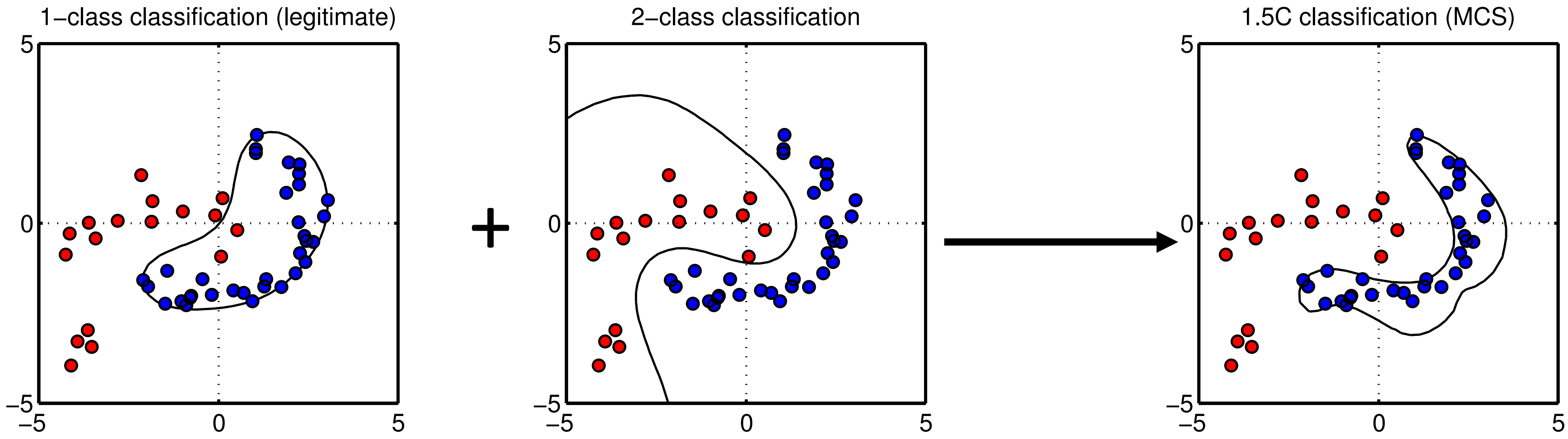
- **Goal**
  maximum-confidence evasion
- **Knowledge**
  perfect
- **Attack strategy**
  compute the minimum modifications
  to the malicious sample so that it falls in the benign area



Building better Class Boundaries

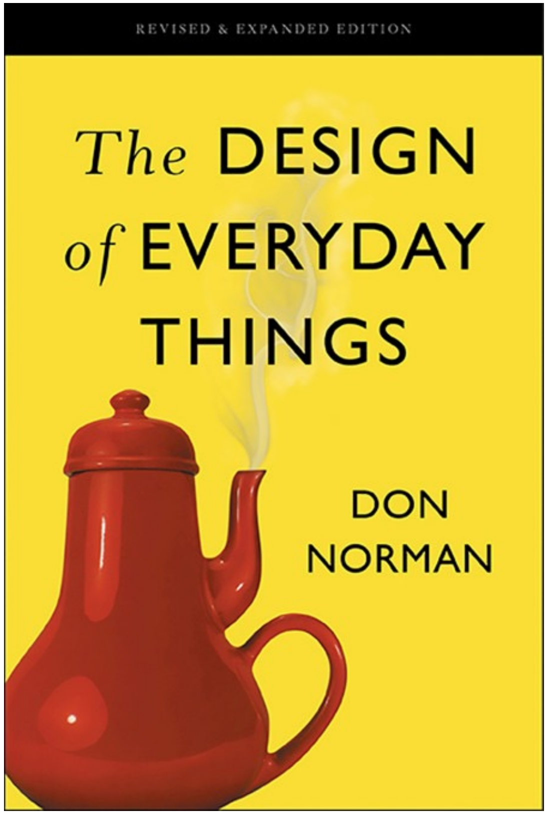
1-class classifiers and 2-class classifiers provide complementary characteristics with respect to evasion attacks

- Different decision boundaries
- Different “no-man’s-land” areas



Prevention and Defence
The roles of usability and awareness

Human-centered design



Coffeepot for Masochists

First edition in 1988, titled
The Psychology of Everyday Things

"Why did you make that error? Didn't you read the manual?"

"Yes, yes, I understand the way it works, but when it comes to practice, I often act automatically and make the error"

Engineers are trained to think logically.

They come to believe that all people must think this way, and they design their machines accordingly.
When people have trouble, the engineers are upset, but often for the wrong reason.

We have to accept human behavior the way it is, not the way we would wish it to be.

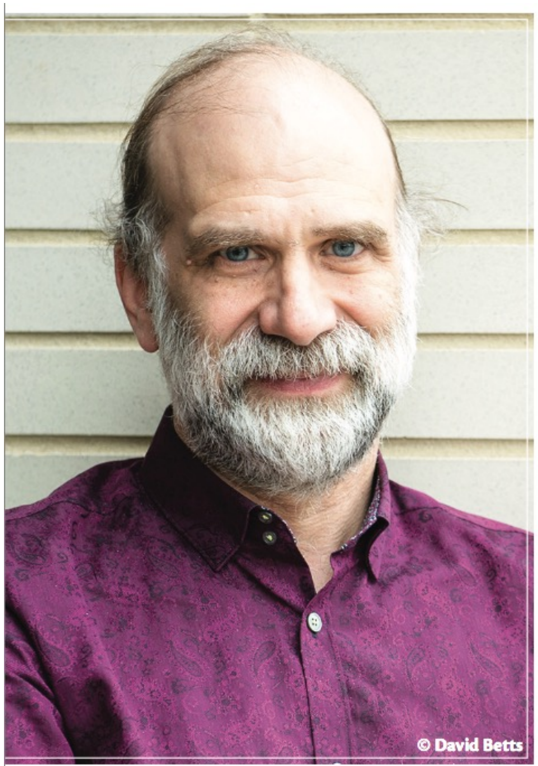
The idea that a person is at fault when something goes wrong is deeply entrenched in society

More and more often the blame is attributed to "human error."

Humans err continually; it is an intrinsic part of our nature.

System design should take this into account

“Stop Trying to Fix the User”



Bruce Schneier
Harvard University

Stop Trying to Fix the User

IEEE Security & Privacy Sept/Oct 2016

Every few years, a researcher replicates a security study by littering USB sticks around an organization’s grounds and waiting to see how many people pick them up and plug them in, causing the autorun function to install innocuous malware on their computers. These studies are great for making security professionals feel superior. The researchers get to demonstrate their security expertise and use the results as “teachable moments” for others. “If only everyone was more security aware and had more security training,” they say, “the Internet would be a much safer place.”

Enough of that. The problem isn’t the users: it’s that we’ve designed our computer systems’ security so badly that we demand the user do all of these counterintuitive things. Why can’t

as a way to bypass the system completely—effectively falling back on the security of their email account.

And finally: phishing links. Users are free to click around the Web until they encounter a link to a phishing website. Then everyone wants to know how to train the user not to click on suspicious links. But you can’t train users not to click on links when you’ve spent the past two decades teaching them that links are there to be clicked.

We must stop trying to fix the user to achieve security. We’ll never get there, and research toward those goals just obscures the real problems. Usable security doesn’t mean “getting people to do what we want.” It means creating security that works, given (or despite) what people do. It means security solutions that

152 Simple Steps to Stay Safe Online:

Security Advice for Non-Tech-Savvy Users

Robert W. Reeder, Iulia Ion, and Sunny Consolvo | Google

IEEE Security and Privacy - September/October 2017

Users often don't follow expert advice for staying secure online, but the reasons for users' noncompliance are only partly understood.

More than 200 security experts were asked for the top three pieces of advice they would give non-tech-savvy users.

The results suggest that, although individual experts give thoughtful, reasonable answers, the expert community as a whole lacks consensus.

Challenges

Security & Safety of AI approaches

- AI needs for **trustworthy data**
- **Data representation** and taxonomy affect the performances of AI to a large extent
- **Interpretability** of AI algorithms enables privacy, security, and safety

AI for Cybersecurity

- **Attacks**
  AI tools used for crafting effective social engineering attacks

- **Defence**
  AI tools used for analysing event data

- AI should be used as an *extension* of human intelligence
  - Machines to perform tasks humans are not good at
  - Machines to *aid* humans perform their tasks
  - Humans to perform tasks machines are not good at

Co-funded by the Erasmus+ Programme of the European Union

EUROPEAN
CYBER
SECURITY
MONTH