Journée ARISTOTE

15 novembre 2018 – Ecole Polytechnique

IA et sécurité :
Nouvelle puissance, nouvelles menaces

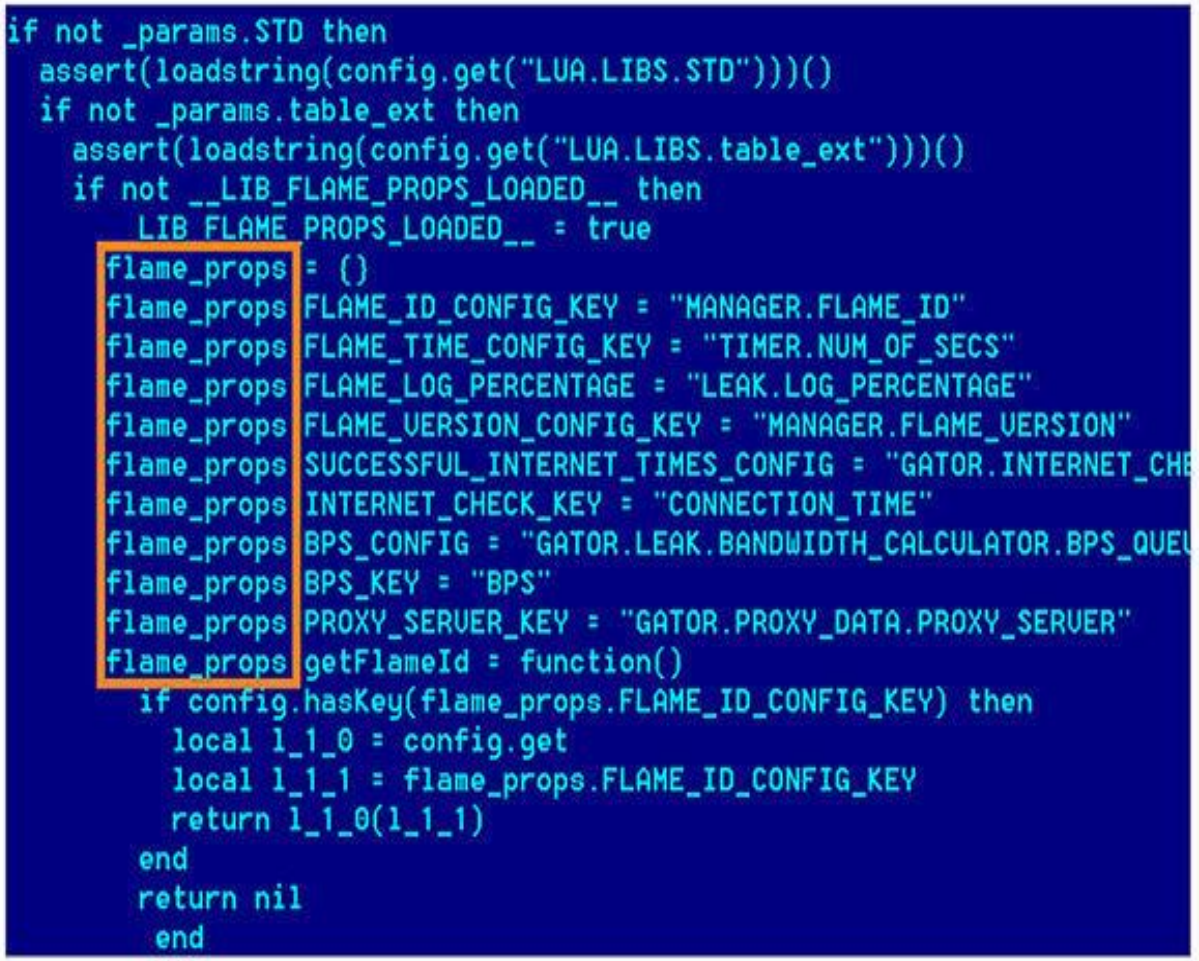
Thierry Berthier
Chercheur associé au CREC Saint-Cyr
Chaire de cybersécurité & cyberdéfense Saint-Cyr
Co-responsable du groupe « Sécurité – IA » du Hub France IA
Membre du conseil scientifique d'ITRUST

http://cyberland.centerblog.net/

I - Panorama de l'insécurité numérique

Creeper : le premier virus de l'histoire de l'informatique, diffusé en 1971 sur le réseau ARPANET ancêtre d'Internet

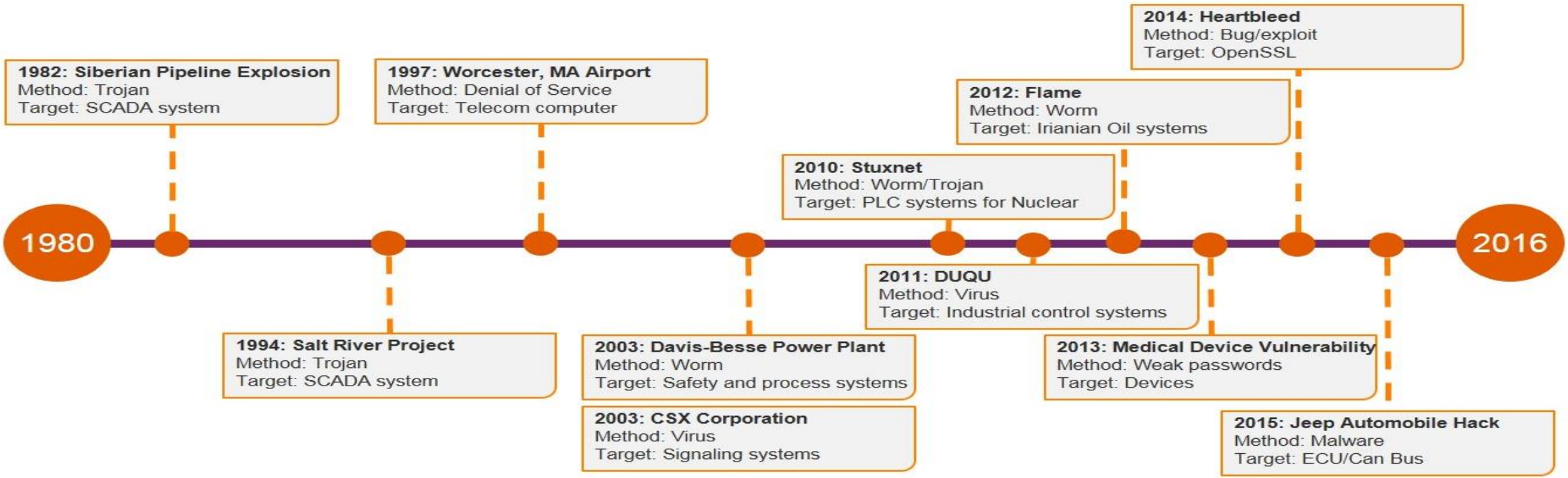
En 1982, Elk Cloner est développé par un programmeur de 15 ans, Rich Skrentaun. Le virus infecte les machines Apple II via une disquette de jeu.

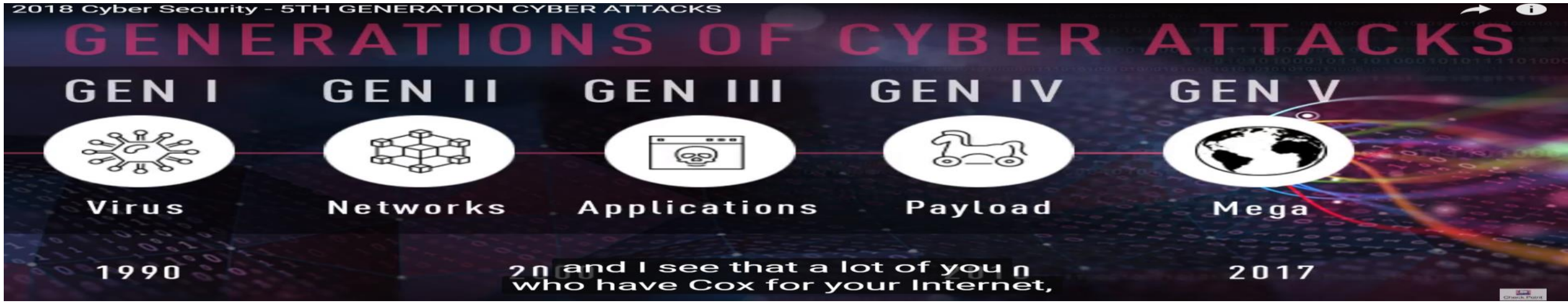


```
if not _params.STD then
  assert(loadstring(config.get("LUA.LIBS.STD")))()
  if not _params.table_ext then
    assert(loadstring(config.get("LUA.LIBS.table_ext")))()
    if not __LIB_FLAME_PROPS_LOADED__ then
        LIB_FLAME_PROPS_LOADED__ = true
      flame_props = {}
      flame_props.FLAME_ID_CONFIG_KEY = "MANAGER.FLAME_ID"
      flame_props.FLAME_TIME_CONFIG_KEY = "TIMER.NUM_OF_SECS"
      flame_props.FLAME_LOG_PERCENTAGE = "LEAK.LOG_PERCENTAGE"
      flame_props.FLAME_VERSION_CONFIG_KEY = "MANAGER.FLAME_VERSION"
      flame_props.SUCCESSFUL_INTERNET_TIMES_CONFIG = "GATOR.INTERNET_CHE
      flame_props.INTERNET_CHECK_KEY = "CONNECTION_TIME"
      flame_props.BPS_CONFIG = "GATOR.LEAK.BANDWIDTH_CALCULATOR.BPS_QUEU
      flame_props.BPS_KEY = "BPS"
      flame_props.PROXY_SERVER_KEY = "GATOR.PROXY_DATA.PROXY_SERVER"
      flame_props.getFlameId = function()
        if config.hasKey(flame_props.FLAME_ID_CONFIG_KEY) then
          local l_1_0 = config.get
          local l_1_1 = flame_props.FLAME_ID_CONFIG_KEY
          return l_1_0(l_1_1)
        end
        return nil
        end
```

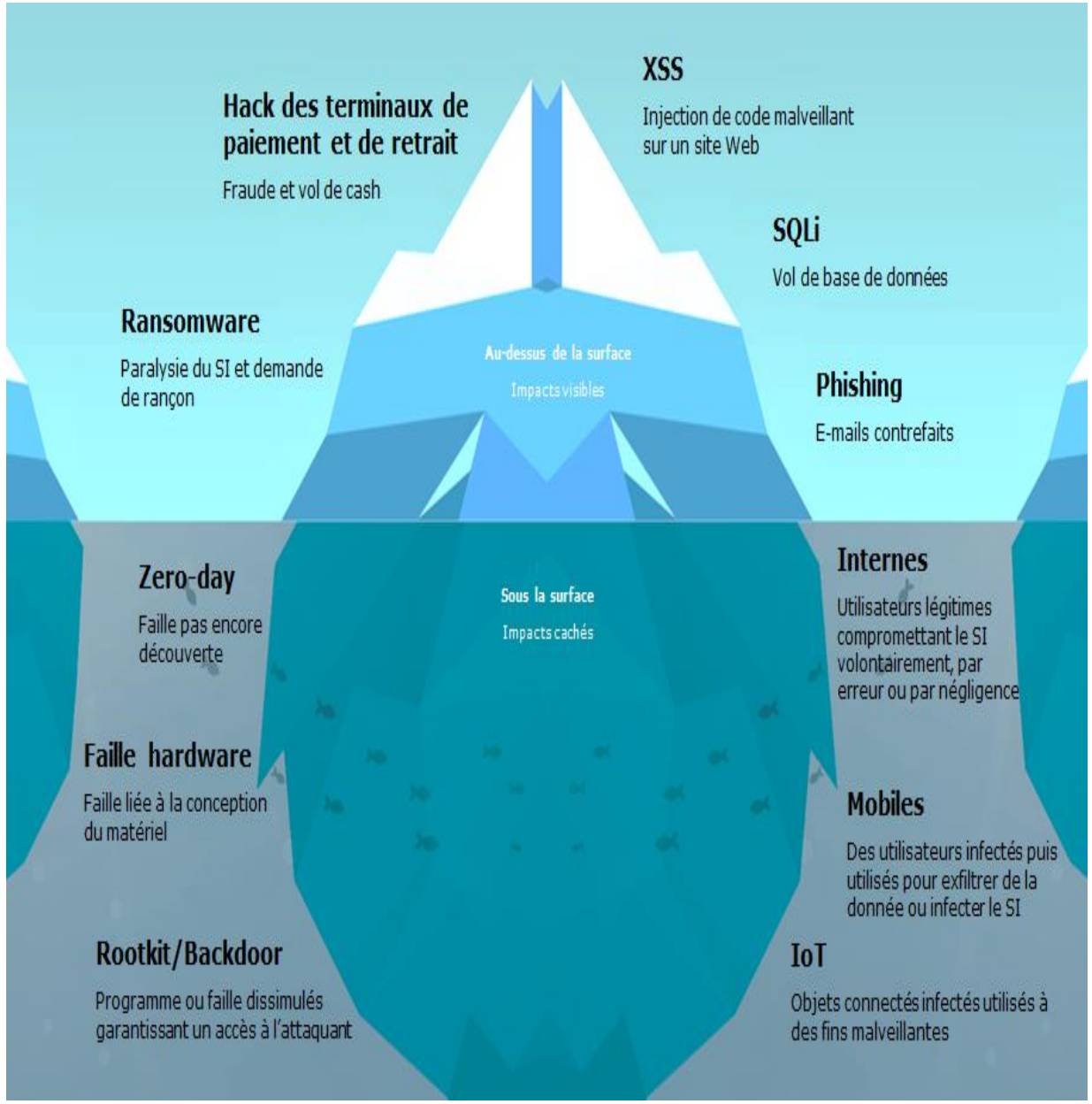
En 2012, le virus de cyberespionnage FLAME se propage dans le monde entier. Il était destiné (quatre ans plus tôt) à l'exfiltration de données sur le programme nucléaire iranien.

Breaches of Industrial Control Systems: 1980-2016



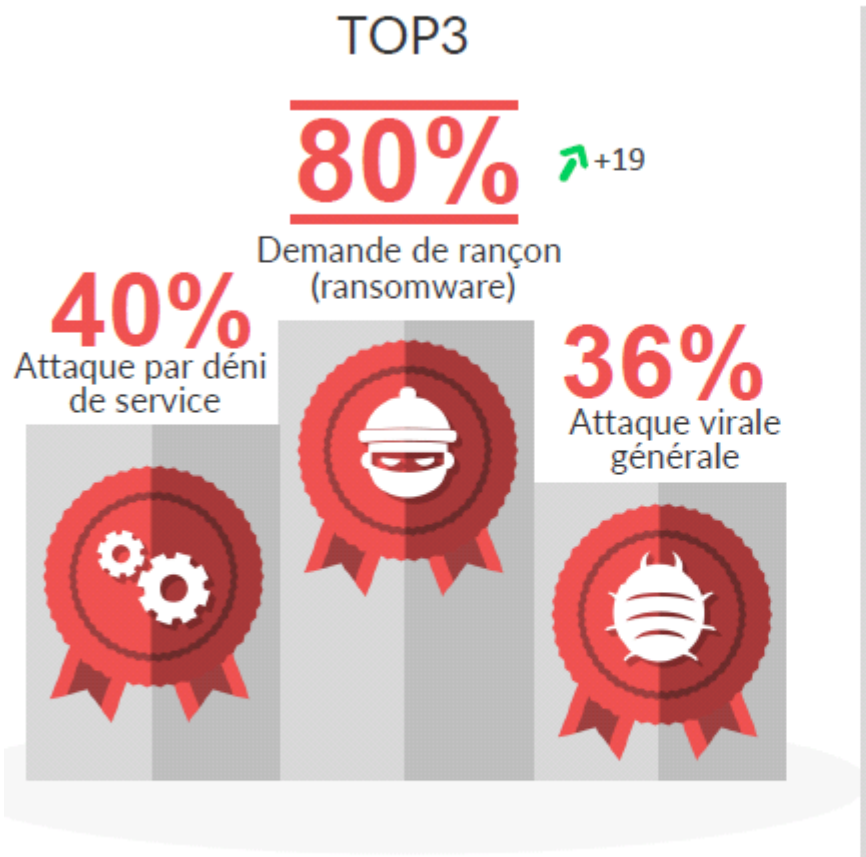


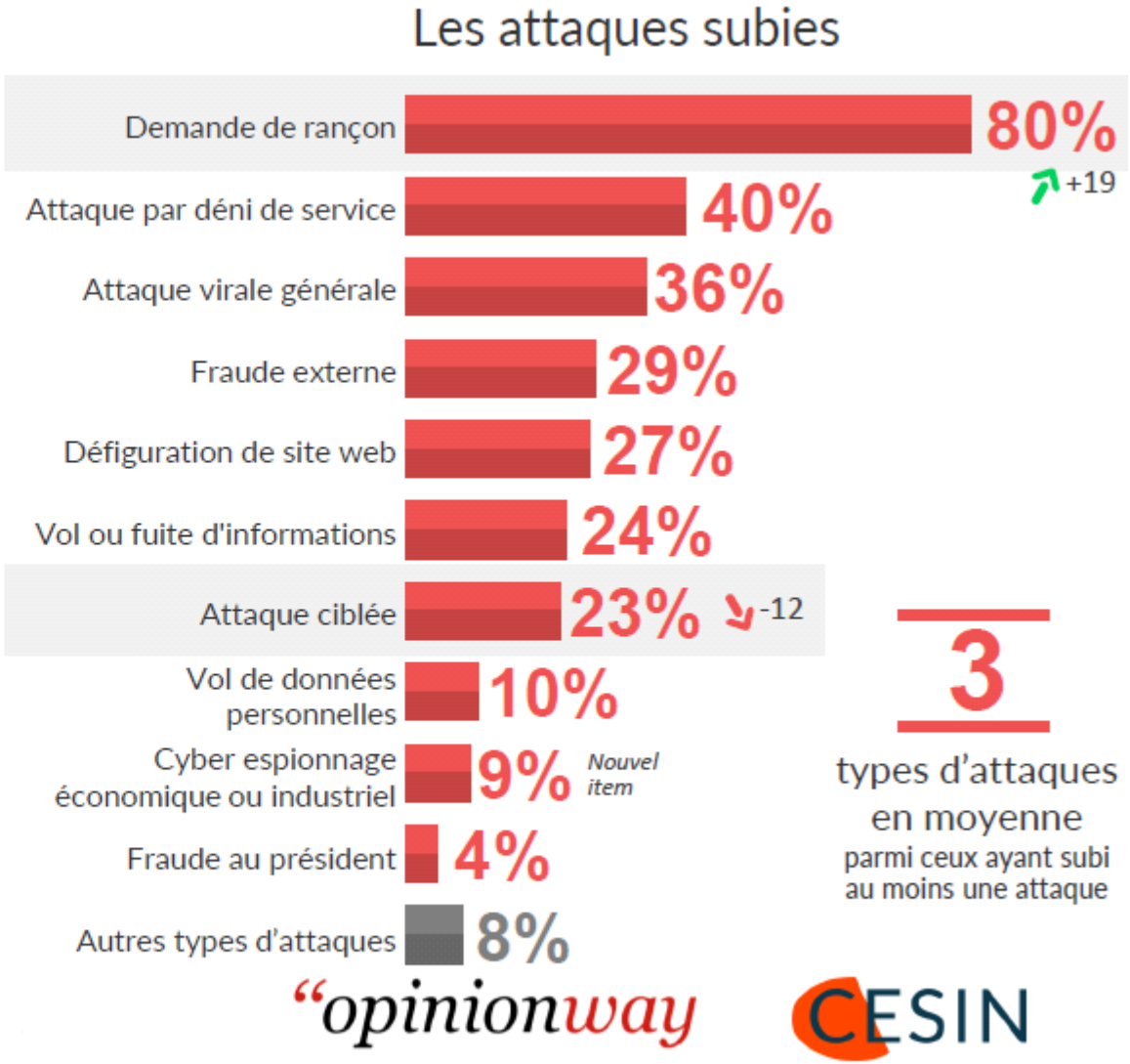
Cartographie des vulnérabilités et des cyberattaques (2016 – 2017)

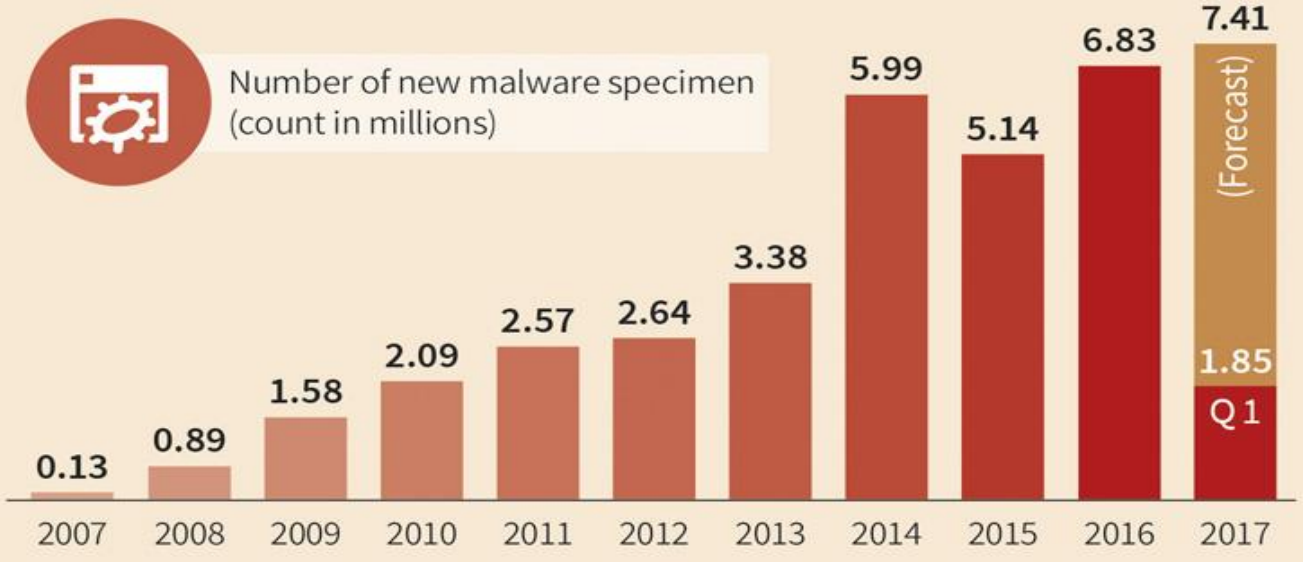


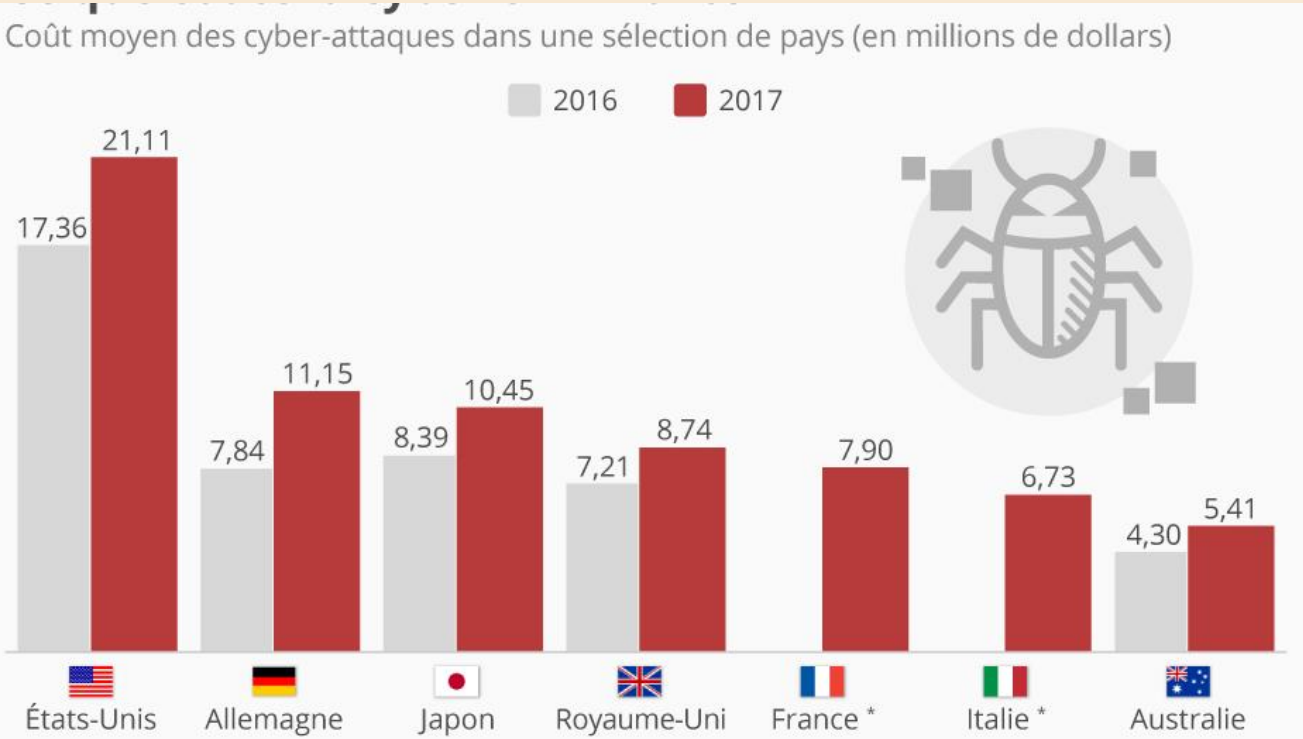
Adware/Publiciel	Logiciel affichant des publicités
Backdoor/Porte dérobée	Logiciel permettant l'accès à distance d'un ordinateur de façon cachée.
Bot	Logiciel automatique qui interagit avec des serveurs.
Exploit	Logiciel permettant d'exploiter une faille de sécurité.
Keyloger/Enregistreur de frappe	Logiciel permettant d'enregistrer les touches frappées sur le clavier.
Ransomware/Rançongiciel	Logiciel qui crypte certaines données du PC, et demande une rançon pour permettre le décryptage.
Rogue	Logiciel se faisant passer pour un antivirus, et indiquant que le PC est gravement infecté. Il se propose de le désinfecter en échange de l'achat d'une licence.
Rootkit	Logiciel permettant de cacher (et de se cacher lui-même) une infection sur un PC.
Spammeur	Logiciel envoyant du spam/pourriel.
Spyware/espiologiciel	Logiciel collectant des informations sur l'utilisateur.
Trojan horse /Cheval de Troie	Logiciel permettant la prise de contrôle à distance d'un PC, il permet souvent l'installation d'une porte dérobée.
Ver/Virus réseau	Logiciel se propageant via un réseau informatique.
Virus	Logiciel conçu pour se propager de PC en PC et s'insérant dans des programmes hôtes.

Typologie des cyberattaques en France – rapport CESIN 2016









www.kaspersky.com

KASPERSKY lab

Kaspersky Lab: Les chiffres majeurs de 2017

Menaces en ligne

Information pour : **2017** | 2016

1 milliard d'attaques malveillantes en ligne : 1 milliard ↑ | 758 millions

15 714 700 Objets uniques malveillants (Scripts, exploits, executable files etc.) détectés par l'antivirus Kaspersky Lab en 2017

22% d'ordinateurs touchés par des programmes publicitaires et leurs composants

88% des attaques sont originaires de 10 pays différents

Ransomware

Plus de **96 000** Modifications de Crypto-ransomware détectés

38 Nouvelles familles découvertes

939 722 Utilisateurs uniques de KSN, attaqués par des crypteurs, incluant...

>240,000 Utilisateurs de nos produits « entreprise ».

Plus de détails les données sont incluses dans le rapport statistique de Kaspersky Security Bulletin

Banking malware

1 126 701 d'appareils ont vu des tentatives d'attaques qui avaient pour but de lancer un malware capable de voler de l'argent via les banques en ligne.

Applications les plus ciblées par les exploits

MS Office: 17.6% ↑ | 13%

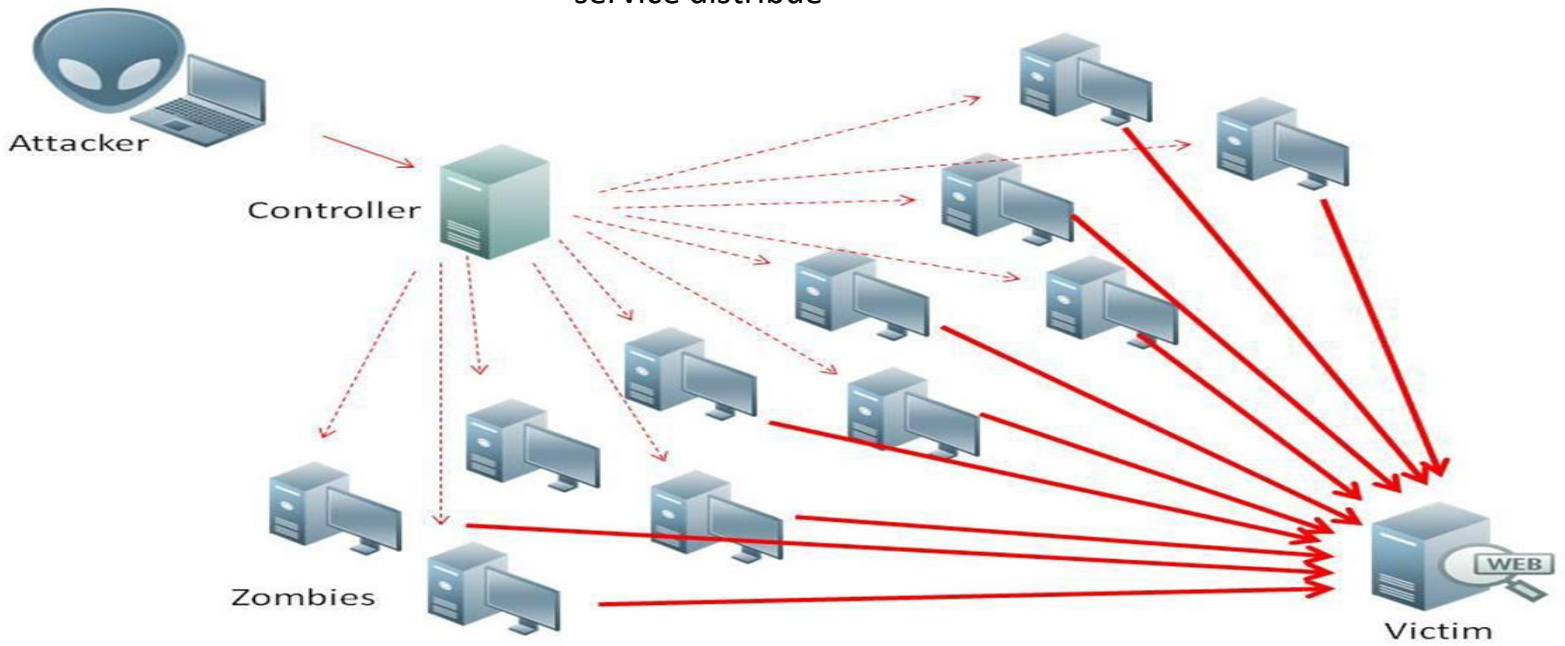
Adobe Flash: 4.5% ↓ | 8%

Statstiques obtenus de Kaspersky Security Network (KSN)

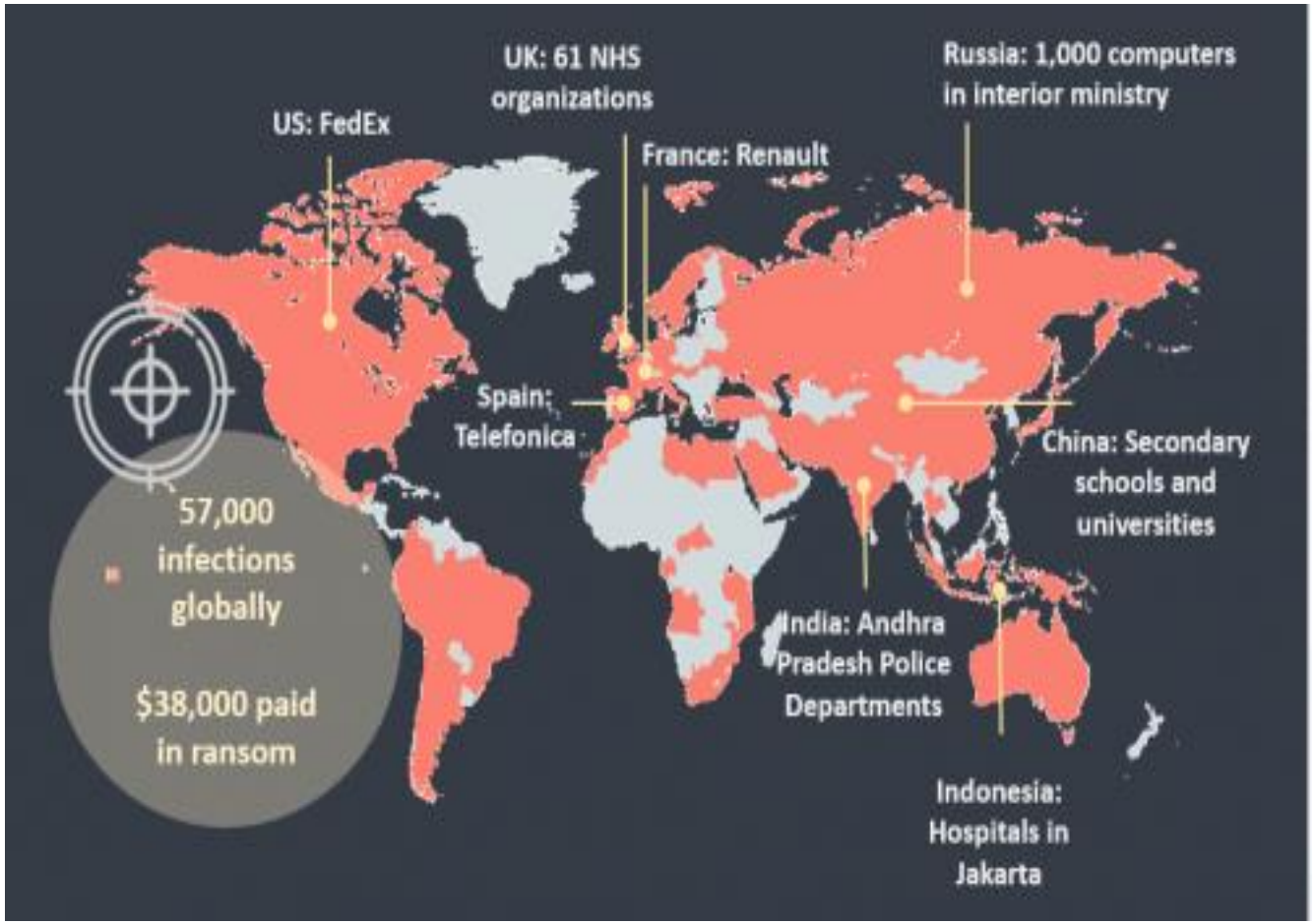
Entre le clavier et le fauteuil… le facteur humain, maillon faible de la chaîne de sécurité

Août 2016 – Cyberattaque sur Ashley Madison – vol de données de 37 millions de membres du site de rencontres extra-conjugales et divorces en cascade

Attaque DDoS : déni de service distribué



WannaCry : l'ère des attaques mondiales



US CyberCommand

China Cyber Army

Des conflits projetés sur le cyberespace

Advanced Threats Are Hard to Find

Cyber Criminals

Nation States

Insider Threats

100%
Valid credentials were used

40
Average # of systems accessed

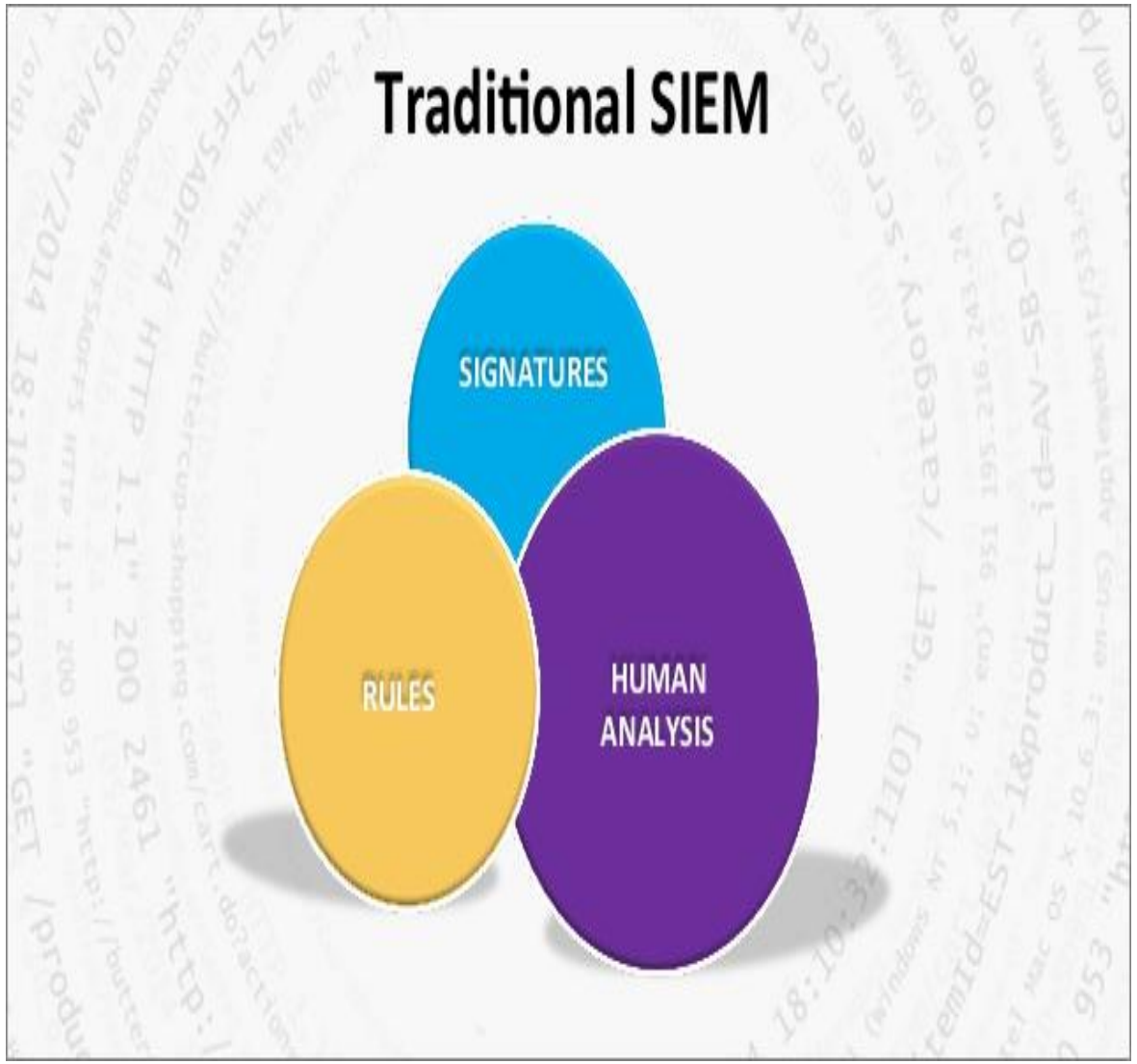
205
Median # of days before detection

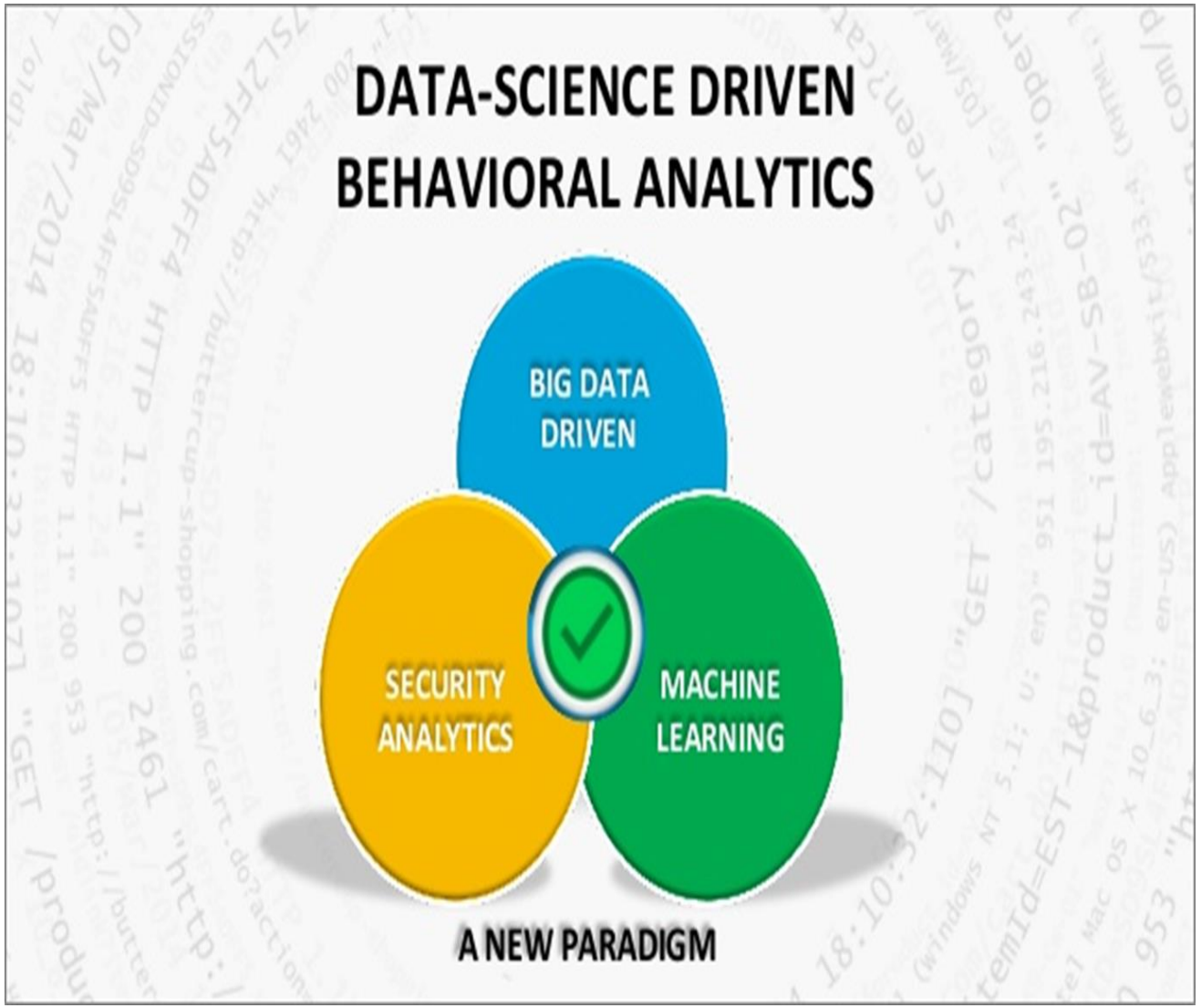
67%
Of victims were notified by external entity

Source: Mandiant M-Trends Report

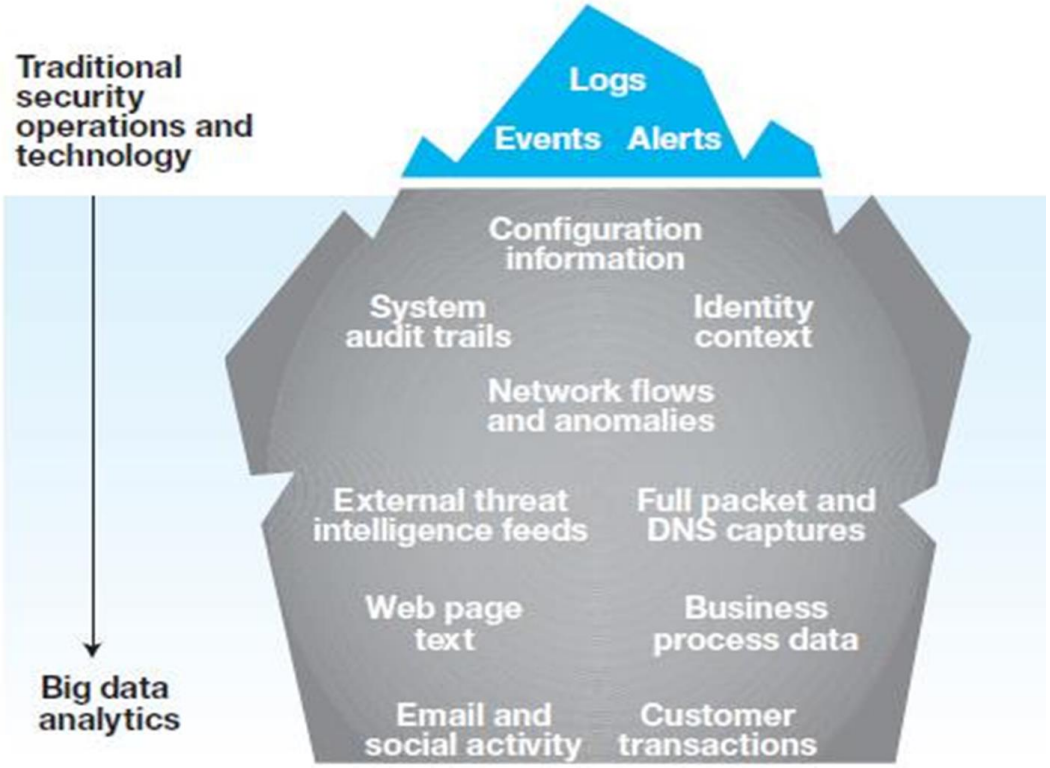
splunk>

II - L'IA au service de la cybersécurité : automatisation de la défense et de l'attaque

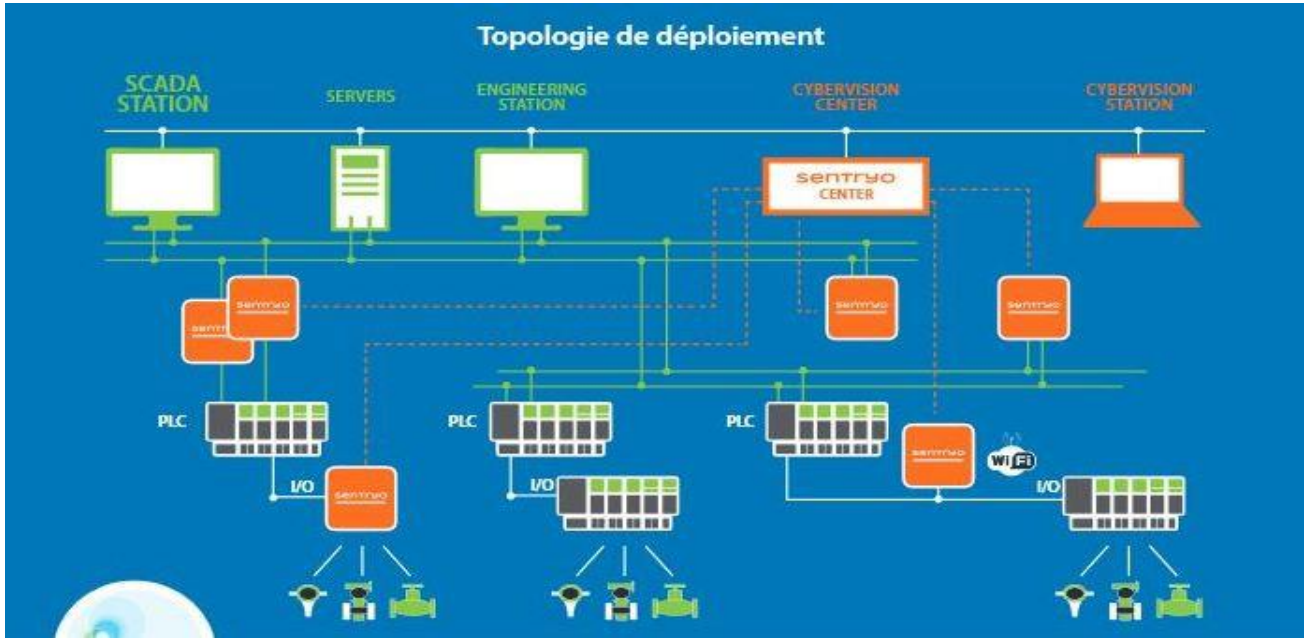


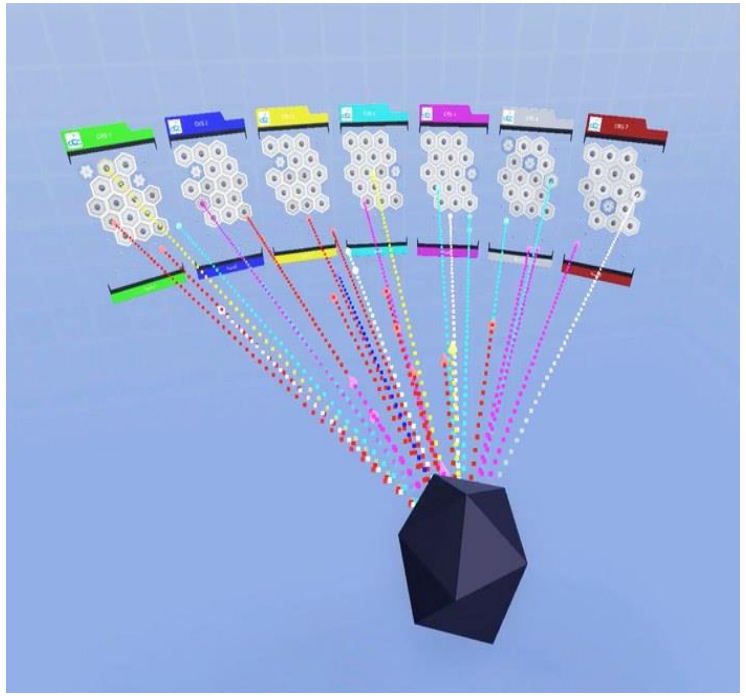


Solution ITrust Reveelium



Solution Sentryo – réseau industriel SCADA





L ’IA va automatiser :

→ de la détection des vulnérabilités,

→ des processus d’attaque,

→ des processus de défense (UBA)

→ de la réponse à incidents,

→ de la sécurité prouvée de certains codes

→ de la sécurité « by design »

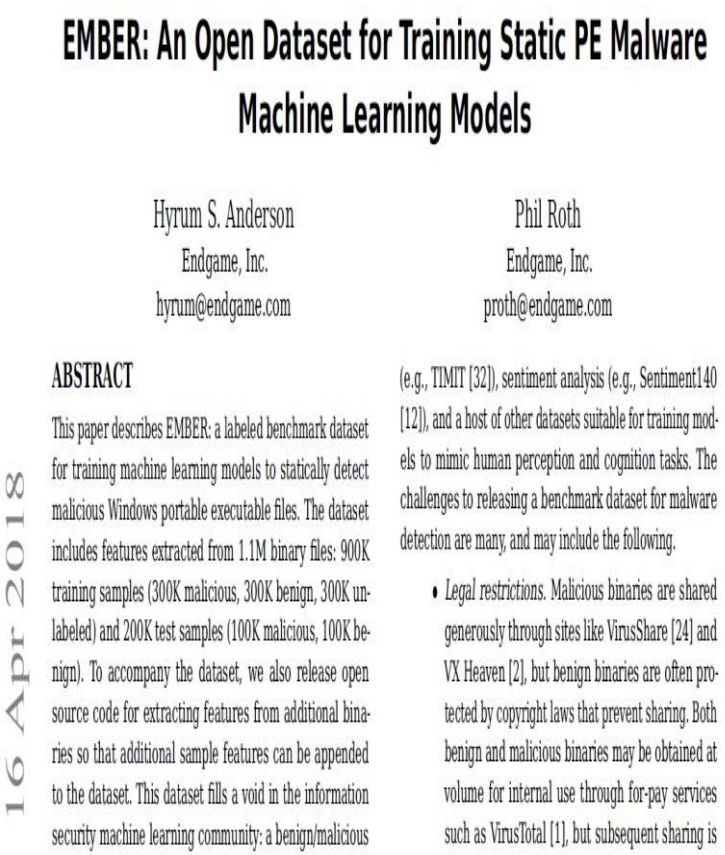
→ De la création d’ADF, architectures de données fictives

DARPA CGC and DEFCON CTF: Automatic Attack and Defense Technique

從DARPA CGC及DEFCON CTF探討自動攻防技術

C.K. Chen

Twitter: Bletchley13



EMBER: An Open Dataset for Training Static PE Malware Machine Learning Models

Hyrum S. Anderson
Endgame, Inc.
hyrum@endgame.com

Phil Roth
Endgame, Inc.
proth@endgame.com

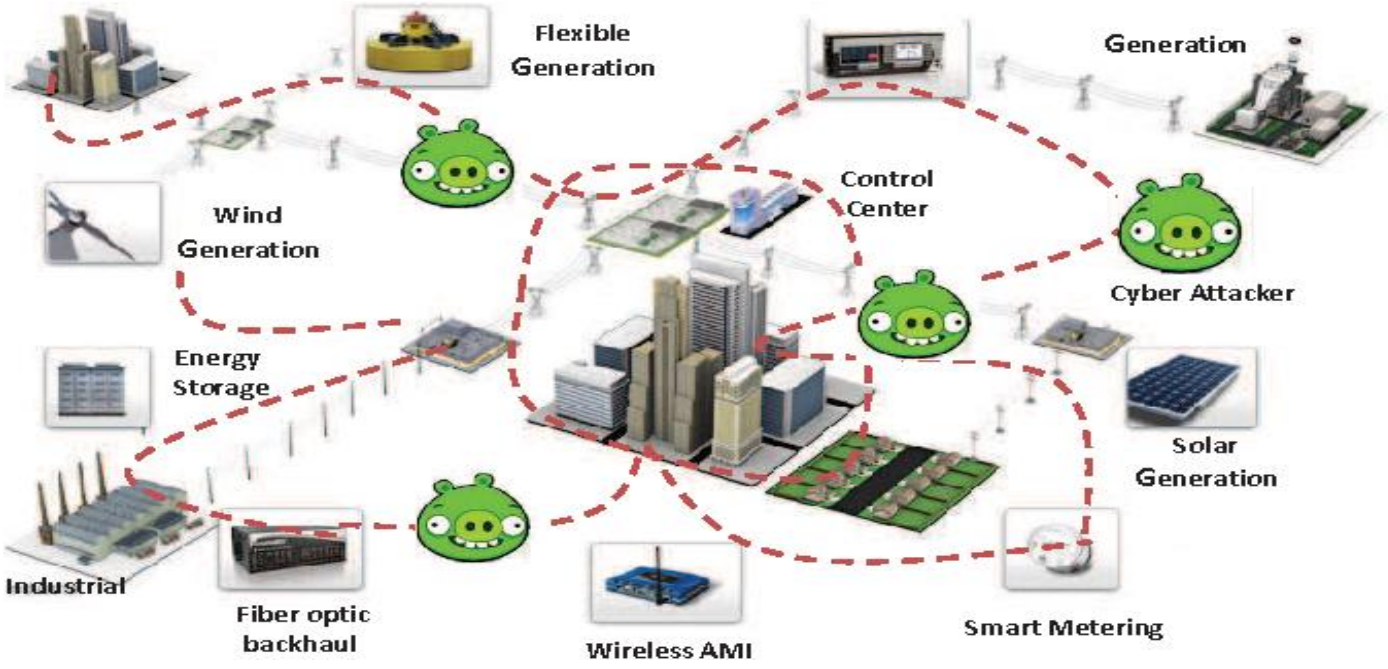
16 Apr 2018

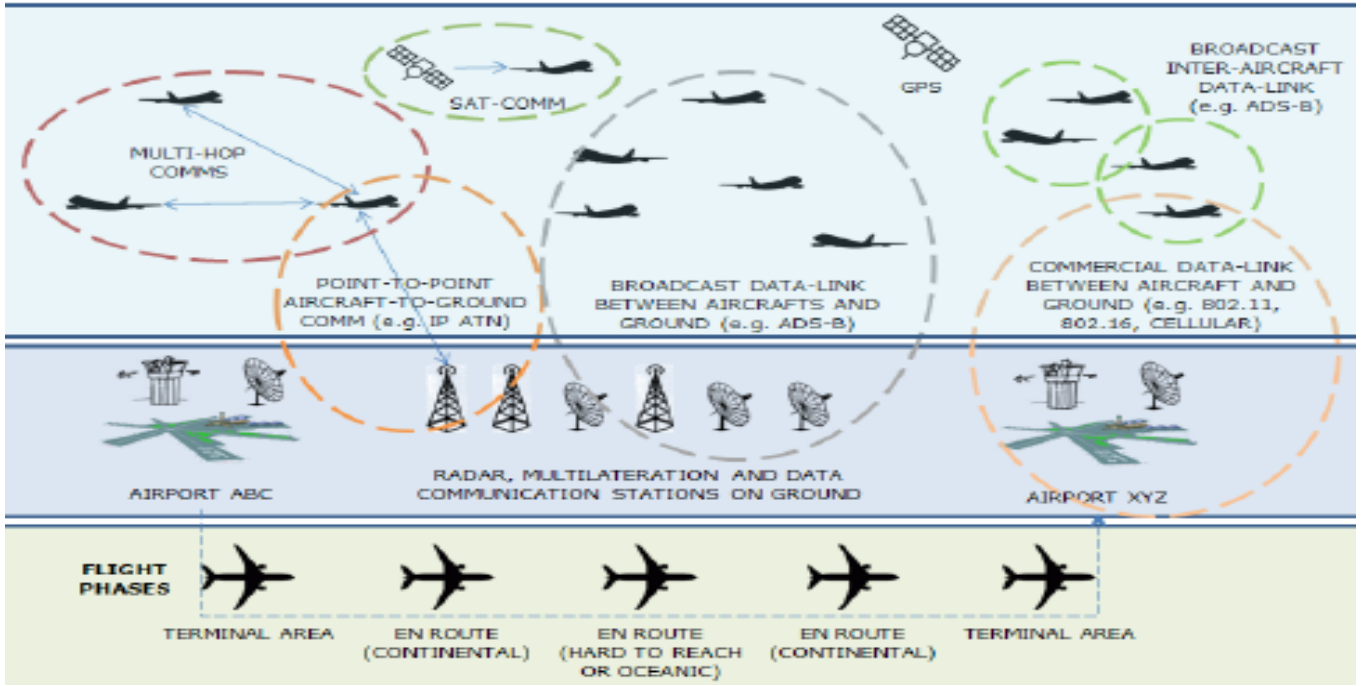
ABSTRACT

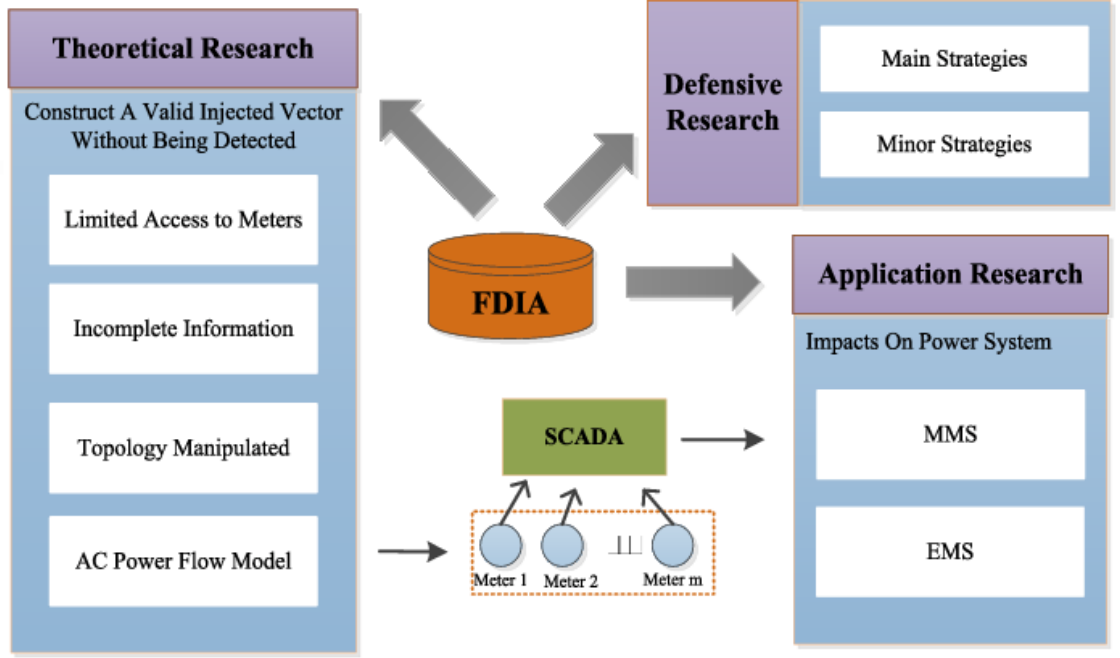
This paper describes EMBER: a labeled benchmark dataset for training machine learning models to statically detect malicious Windows portable executable files. The dataset includes features extracted from 1.1M binary files: 900K training samples (300K malicious, 300K benign, 300K unlabeled) and 200K test samples (100K malicious, 100K benign). To accompany the dataset, we also release open source code for extracting features from additional binaries so that additional sample features can be appended to the dataset. This dataset fills a void in the information security machine learning community: a benign/malicious

(e.g., TIMIT [32]), sentiment analysis (e.g., Sentiment140 [12]), and a host of other datasets suitable for training models to mimic human perception and cognition tasks. The challenges to releasing a benchmark dataset for malware detection are many, and may include the following.

- *Legal restrictions.* Malicious binaries are shared generously through sites like VirusShare [24] and VX Heaven [2], but benign binaries are often protected by copyright laws that prevent sharing. Both benign and malicious binaries may be obtained at volume for internal use through for-pay services such as VirusTotal [1], but subsequent sharing is







Exemple 1 : FDIA sur des composantes ML de Smart Grids

Exemple 2 : FDIA sur des composantes ML de contrôle aérien ADS-B

Les nouveaux risques liés à l'IA: ATTAQUES FDIA sur des composantes Machine Learning – False Data Injection Attacks

Les réseaux résilients CISCO

cisco

Threat Detection and Mitigation for IoT Systems using Self Learning Networks (SLN)

Self Learning Networks
A true paradigm shift

Signature-based (Firewalls): traffic is normal ***unless*** matching known characteristics

SLN

Dynamic Learning of anomalies (SLN): Outperform conventional algorithms in presence of uncertainty, when complexity is too large (scale) and when adaptation is required.

This is a key requirement in IoT/IoE. We need predictive models for large scale networks to address:

- High performance and high resiliency
- Detection of disruptive subtle DDoS attacks
- New threats

Threat Detection
Anomaly Detection: A paradigm change

Traditional Anomaly Detection Systems
Focus on Detection (*wrong*)
Core challenge is *not* Detection itself but Precision (avoid False Positive / Irrelevant alarms)

SLN Approach
Efficient detection *and* Precision
Make the Network learn from its own mistakes and eliminate False Positive !
Not a feature but an *Architecture*
Dynamic Learning of anomalies (SLN): mathematical models built on-line and deviation from these (constantly adapted) models lead to detection of anomalies

Threat Detection
Anomaly Detection: A paradigm change

Traditional Anomaly Detection Systems
Focus on Detection (*wrong*)
Core challenge is *not* Detection itself but Precision (avoid False Positive / Irrelevant alarms)

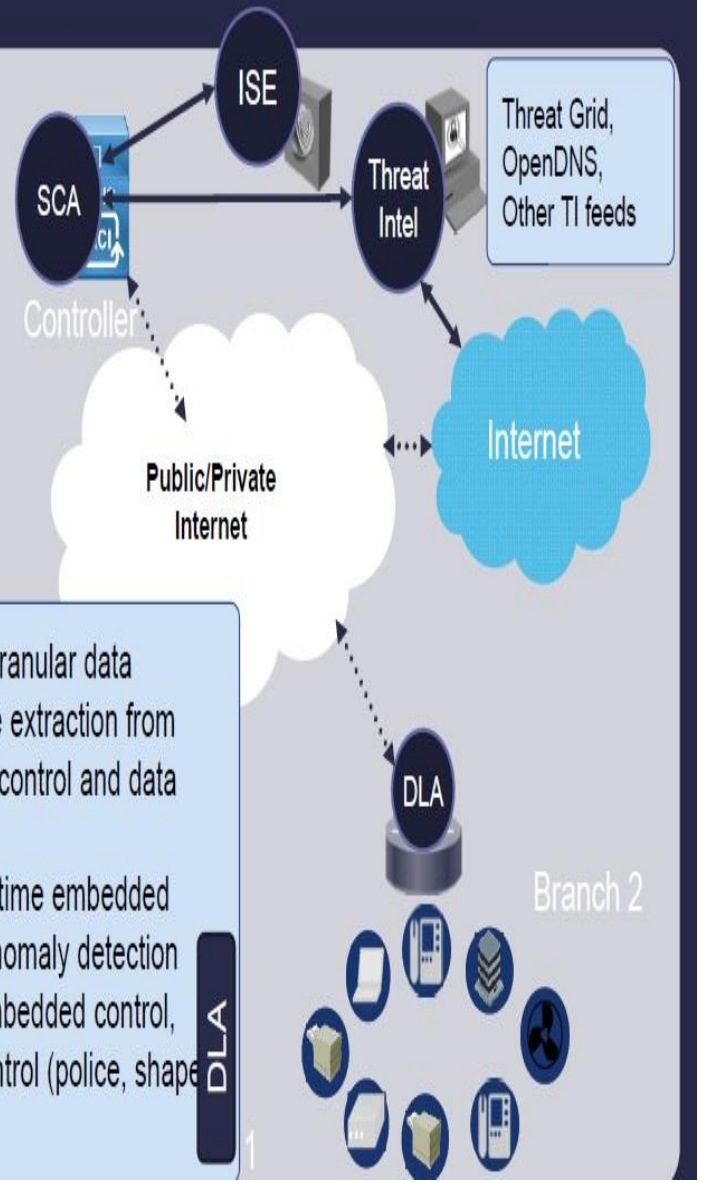
SLN Approach
Efficient detection *and* Precision
Make the Network learn from its own mistakes and eliminate False Positive !
Not a feature but an *Architecture*
Dynamic Learning of anomalies (SLN): mathematical models built on-line and deviation from these (constantly adapted) models lead to detection of anomalies

Des systèmes résilients « by design »

SLN Architecture

- Orchestration of Distributed Learning Agents (DLAs)
- Advanced Visualization of anomalies
- Centralized policy for mitigation
- Interaction with other security components such as ISE and Threat Intelligence Feeds
- North bound API to SIEM/Database
- Evaluation of anomaly relevancy

SCA

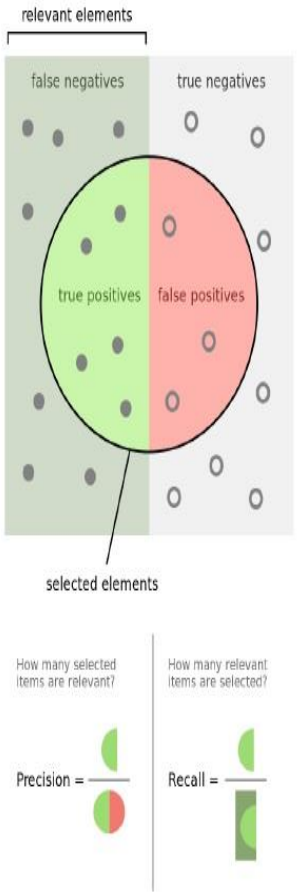


- **Sensing (knowledge)**: granular data collection with knowledge extraction from NetFlow but also DPI on control and data plane & local states
- **Machine Learning**: real-time embedded behavioral modeling & anomaly detection
- **Control**: autonomous embedded control, advanced networking control (police, shape, recoloring, redirect, ...)

DLA

Discussing Recall, Precision, FP, ...

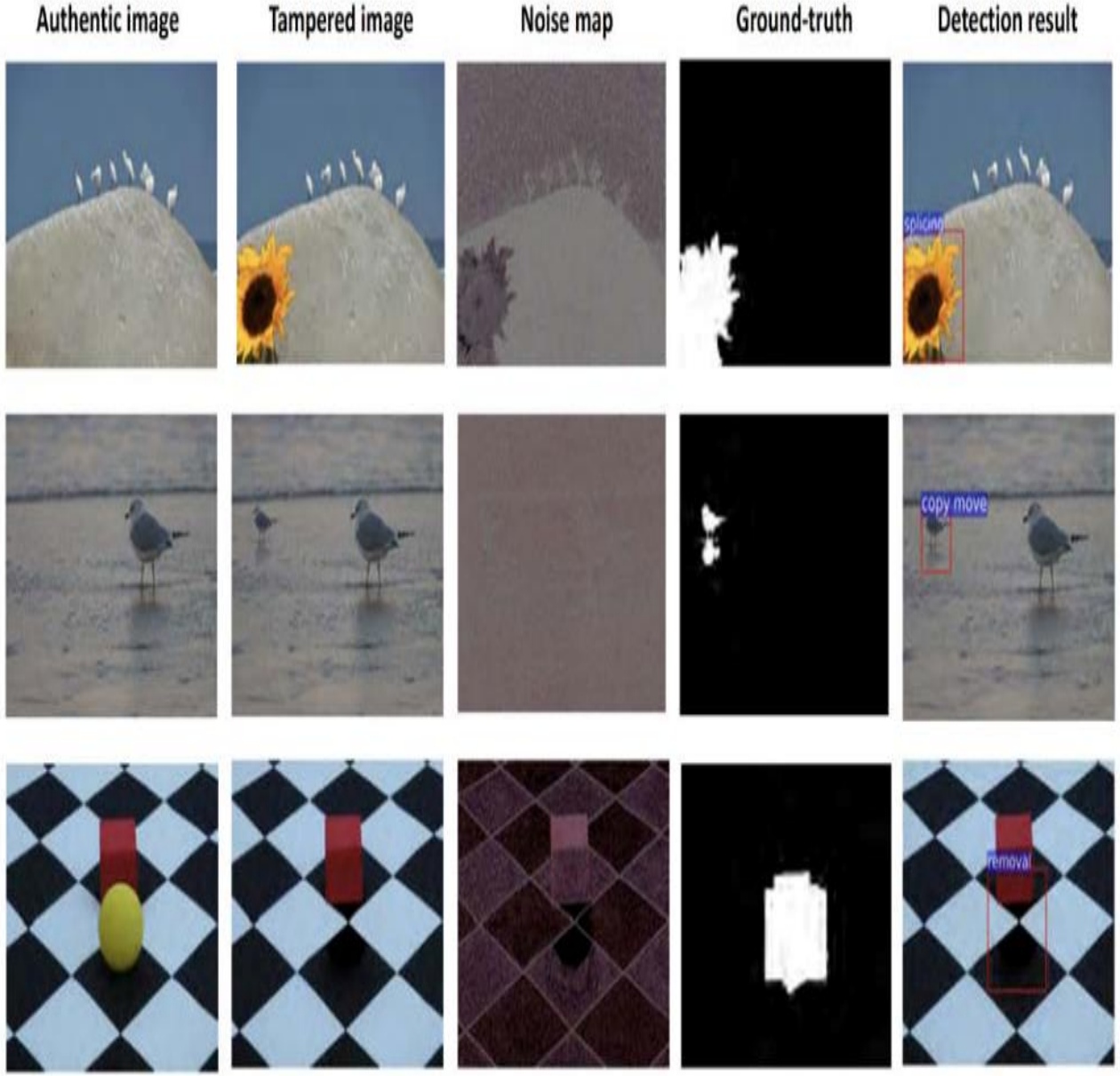
- Few simple notions required when discussing Machine Learning: False Positive (FP), True Positive (TP), False Negative (FN), True Negative (TN), Recall and Precision.
- Take a Classifier C trained to detect if an event E is relevant (Like) or not (irrelevant).
  - TP: E is classified as relevant and is indeed an relevant
  - FP: E is classified as relevant and is in fact irrelevant (noise)
  - TN: E is classified as irrelevant and is indeed irrelevant
  - FN: E is classified as irrelevant and is in fact an relevant
- Recall = TP / (TP + FN) (notion of sensitivity)
- Precision = TP / (TP + FP) (positive predictive value)
- Accuracy ACC = (TP + TN) / (TP + TN + FP + FN),
- *Example: if a classifier that is trained to detect dogs in a picture detects 15 dogs, only 10 of them are dogs, and there are 20 dogs in the picture then the Precision = 10/15 = 0.66 and Recall = 10/20 = 0.50*



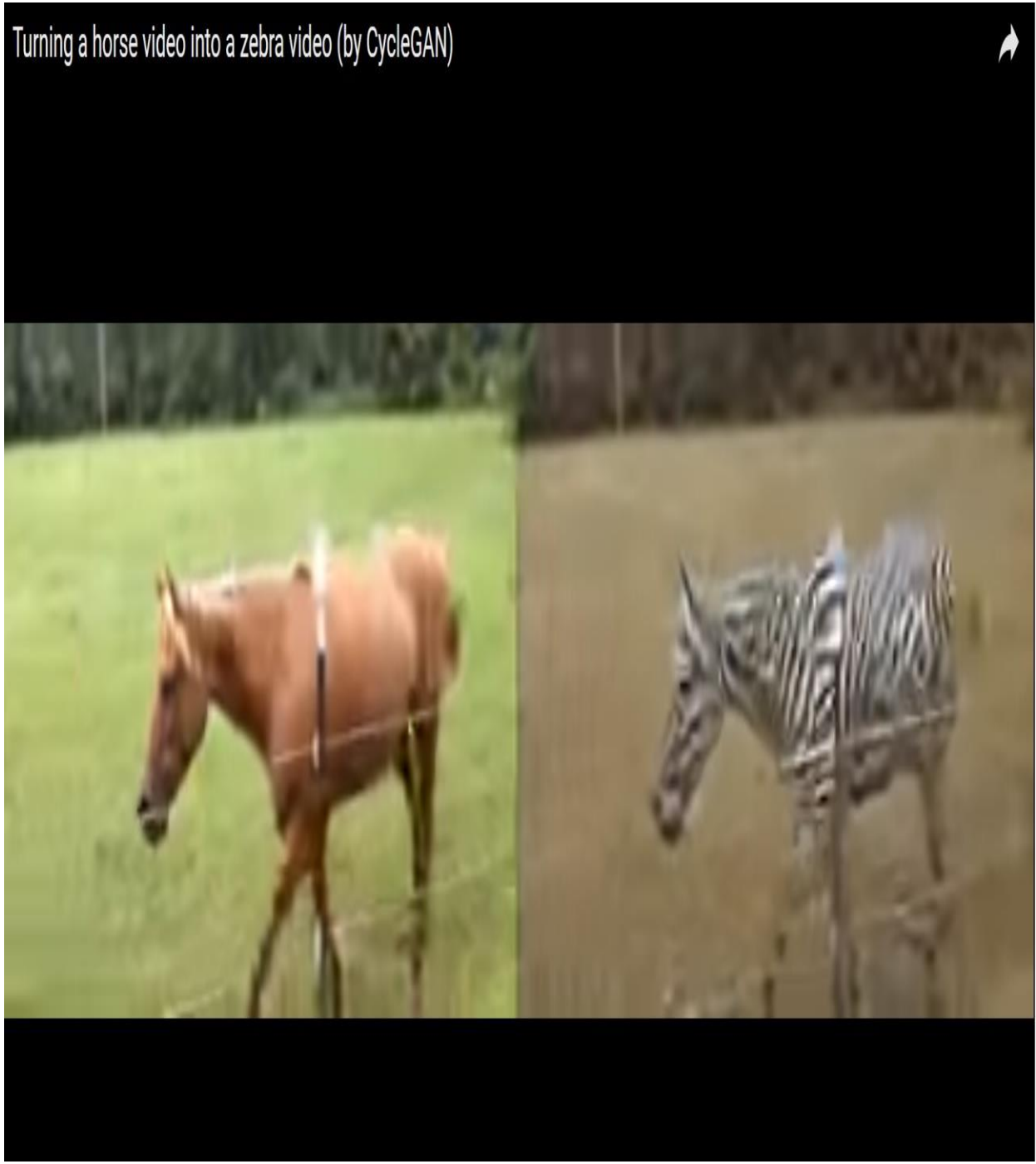
III - L'IA génératrice d'Architectures de Données Fictives Immersives (ADFI)

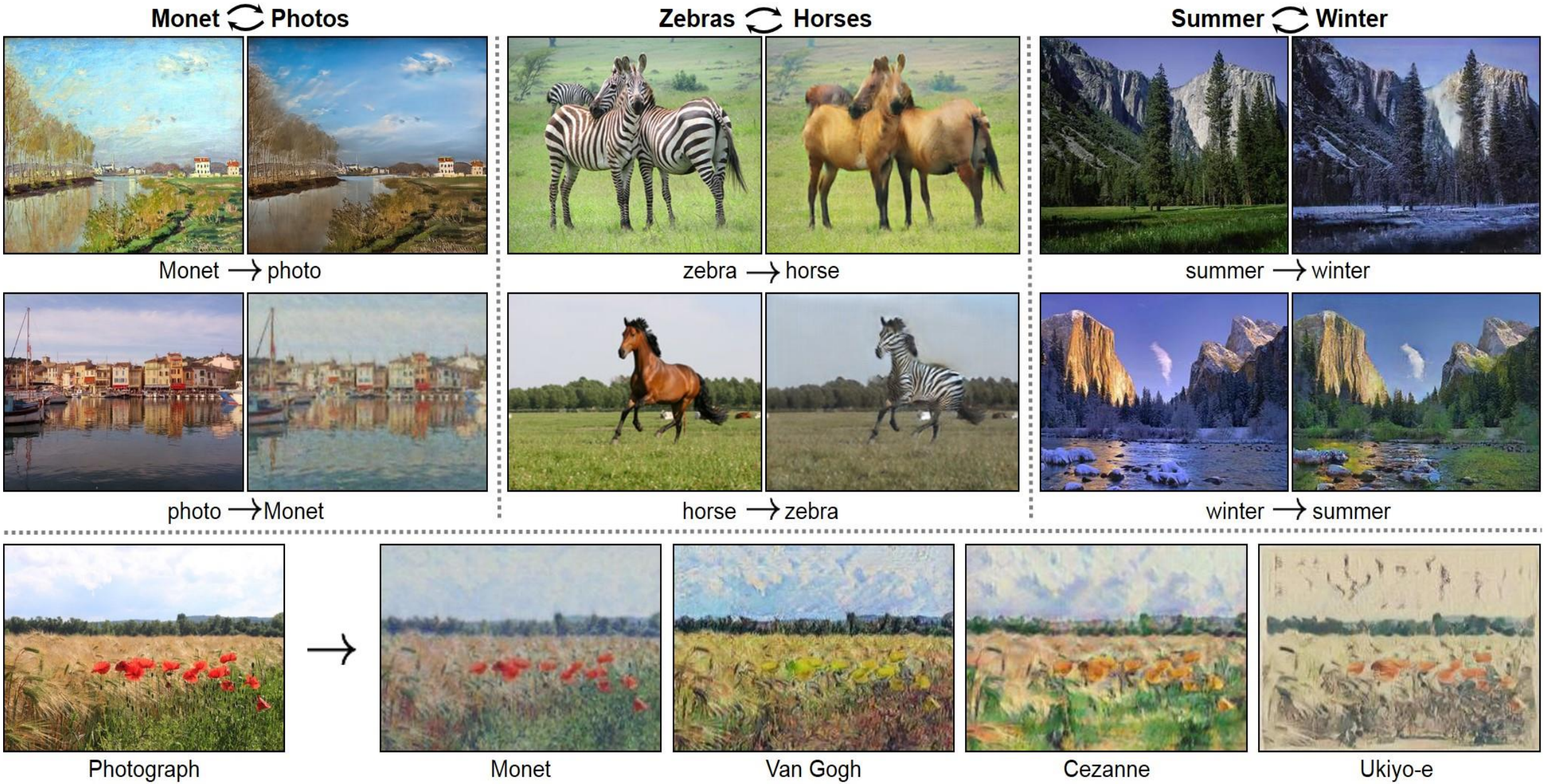
Nous serons bientôt confrontés à des Architectures de Données Fictives (ADF) immersives, sophistiquées, crédibles qui s'appuieront sur nos biais cognitifs, nos fragilités émotionnelles et biologiques pour nous tromper et pour exploiter pleinement le « facteur humain » en attaque.

Dans la matrice ? - hors la matrice ?



An example of authentic images, manipulated images, the RGB and noise streams used to detect manipulation, and the results of AI analysis. Source: the NC2016 dataset




Monet ⇄ Photos
Zebras ⇄ Horses
Summer ⇄ Winter
Monet → photo
zebra → horse
summer → winter
photo → Monet
horse → zebra
winter → summer
Photograph
Monet
Van Gogh
Cezanne
Ukiyo-e

Face2Face: Real-time Face Capture and Reenactment of RGB Videos (CVPR 2016 Oral)

Real-time Reenactment

Reenactment Result

2:54 / 6:35

Synthesizing Obama: Learning Lip Sync from Audio

Synthesizing Obama: Learning Lip Sync from Audio

Supasorn Suwajanakorn
Steven M. Seitz
Ira Kemelmacher-Shlizerman
University of Washington

SIGGRAPH 2017

http://grail.cs.washington.edu/projects/AudioToObama/

Original Video for Input Speech

Our Result

LIPING/ SHUTTERSTOCK.COM

How Fake Data Can Help the Pentagon Track Rogue Weapons

This software trains machine-learning tools to spot terrorists stockpiling weapons.

By Jack Corrigan

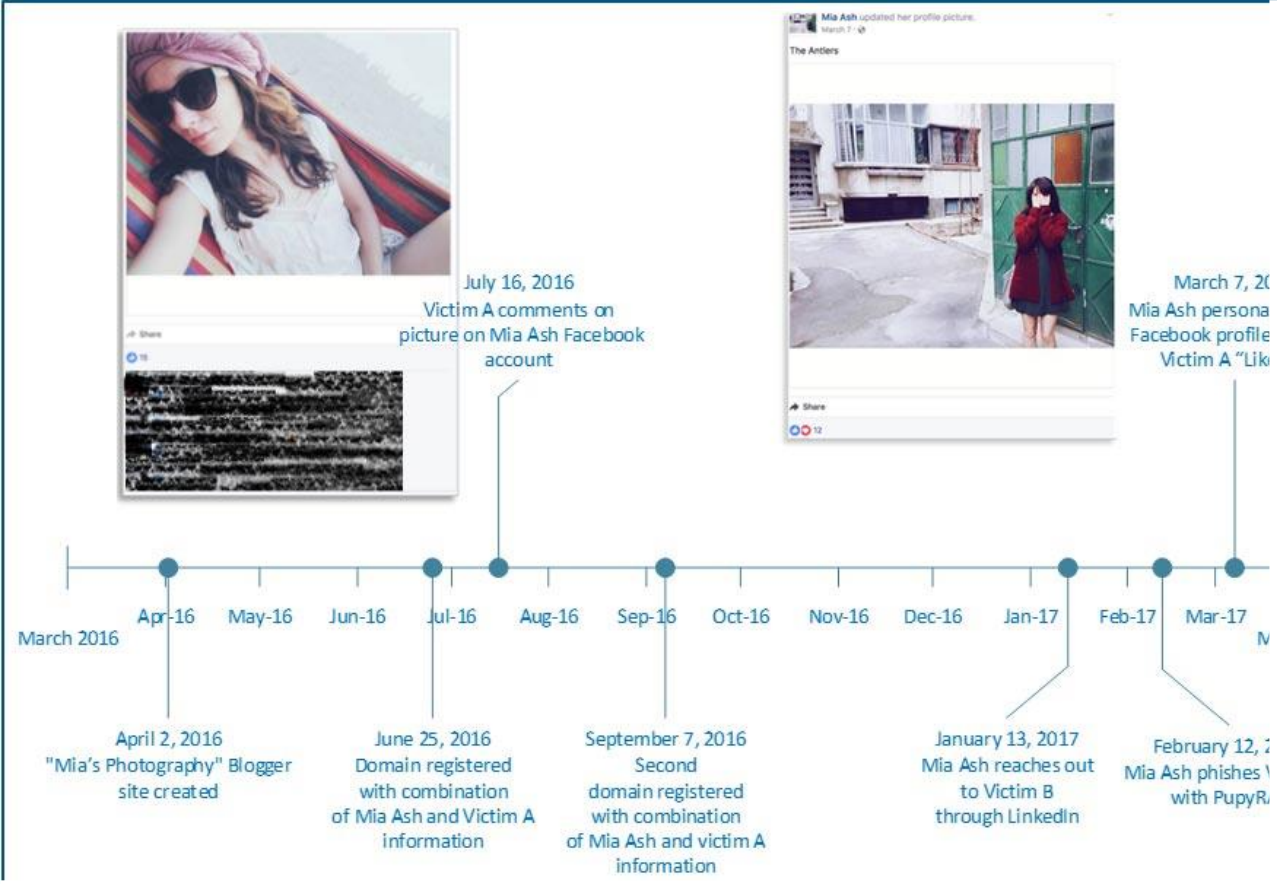
The Pentagon is investing in software that uses big data to help intelligence officers keep terrorists from getting their hands on biological, chemical and nuclear weapons.

The Air Force Research Laboratory in January announced a $4.6 million contract with the software company IvySys to model different ways state and non-state actors could obtain and deploy "weapons of mass terror" around the world.

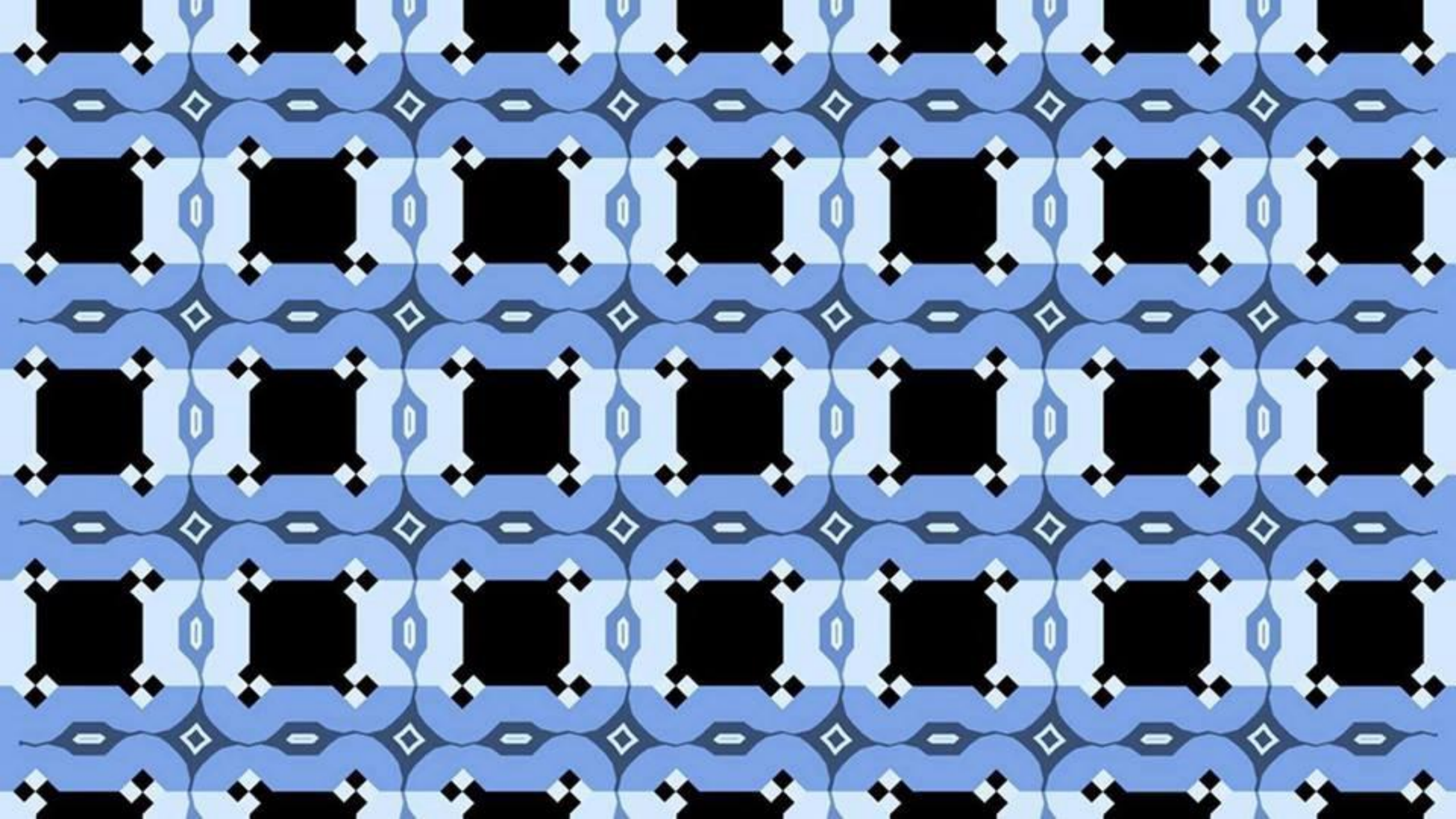
The contract supports an ongoing effort by the Defense Advanced Research Projects Agency to build tools to spot groups who are potentially stockpiling materials for such weapons.

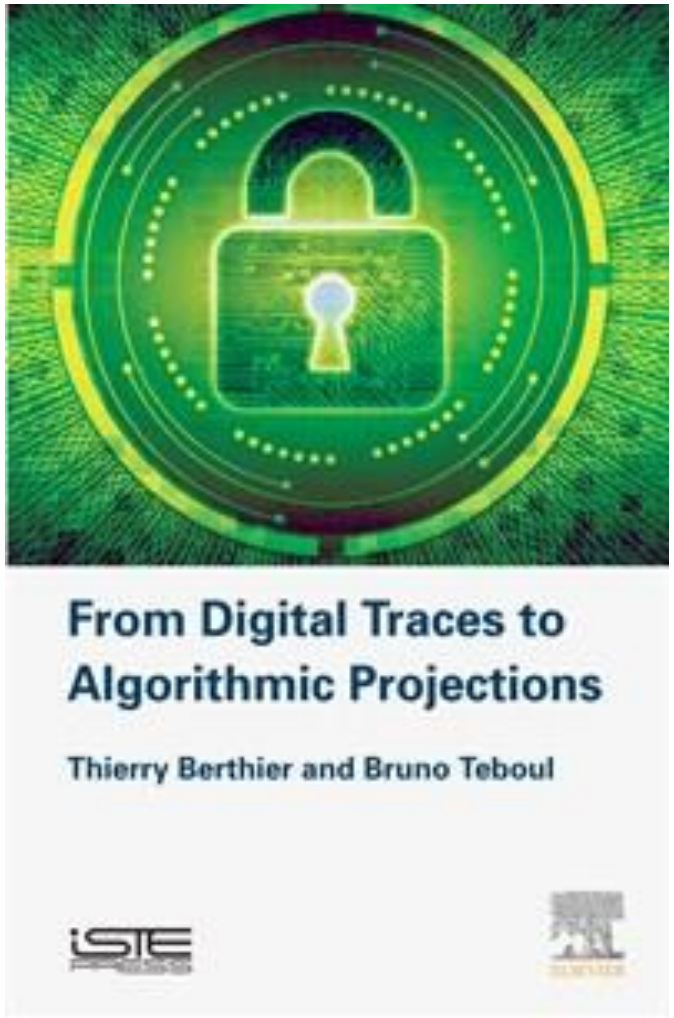
"Reports of chemical weapons use around the world raises serious concerns about non-state actors' access to weapons of mass terror and reinforces fears of a possible terrorist attack with chemical, biological, radiological, or nuclear weapons in the West," DARPA and IvySys said in a statement. "Today's terrorist networks move operatives, money and material across borders and through the crevices of the global economy, making tracking such adversaries a daunting challenge."

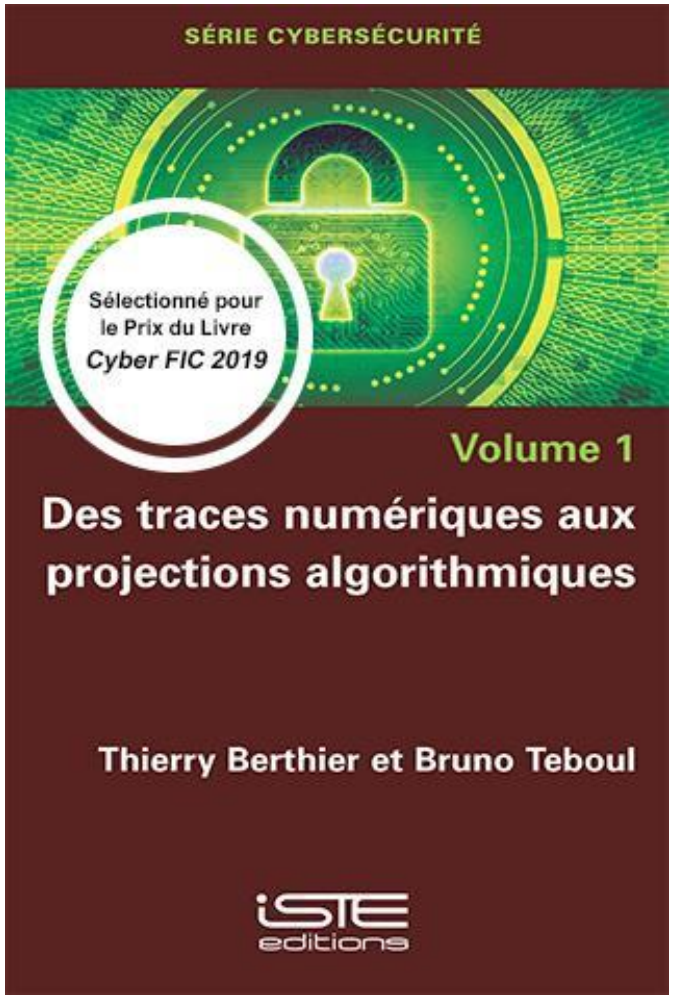
The technology would generate fictional but realistic datasets of bank transactions, emails and inventory transfers, and embed them with indicators of suspicious activities, like a shipment of toxic chemicals getting



Opération COBALT 2016 – 2017 (IRAN ?)







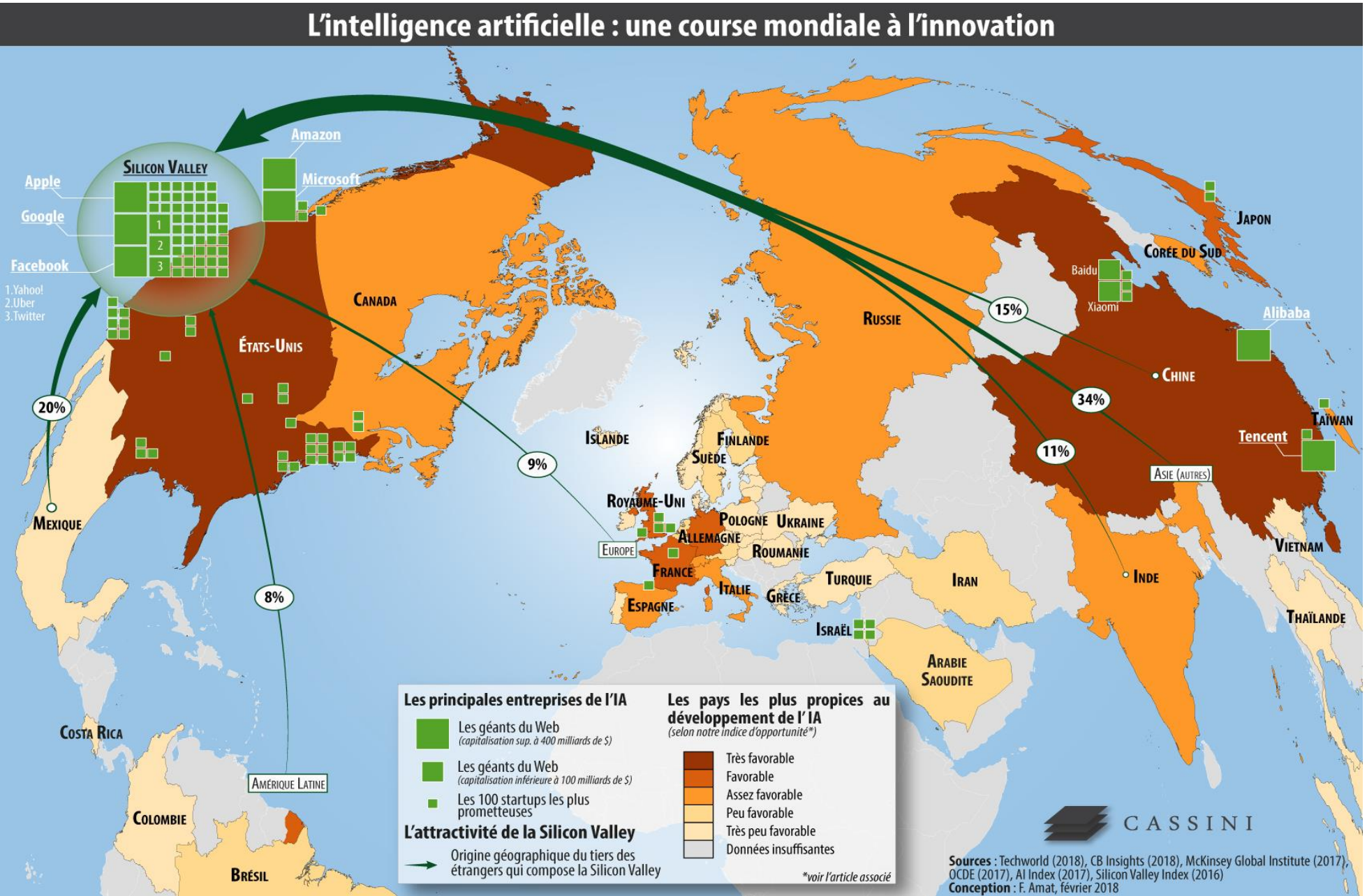
Pour aller plus loin sur les ADFI

https://www.elsevier.com/books/from-digital-traces-to-algorithmic-projections/berthier/978-1-78548-270-0

http://www.iste.co.uk/book.php?id=1372

https://iste-editions.fr/products/des-traces-numeriques-aux-projections-algorithmiques

IV - Cyberstratégie, sécurité et géopolitique de l'IA



What the CIA's Tech Director Wants from AI

Dawn Meyerriecks says staying ahead of Russia and China isn't as hard as getting U.S. leaders to listen to their own artificial intelligence analysis

By Patrick Tucker

Opposite: High Performance Computing and Storage Complex II (HRSK-II) during the official opening of the new data center of the Lehmann Center (LZR) of the Dresden University of Technology in Dresden, eastern Germany, Wednesday, May 13, 2015. / AP/Jens Meyer

Should the U.S. fear growing Russian progress in artificial intelligence? Last week, Vladimir Putin told students, "Whoever becomes the leader in this sphere will become the ruler of the world." That caught the interest of noted AI phobe / profiteer Elon Musk who tweeted, simply and ominously: "It begins..."

But the CIA's head of technology development has a different take. Dawn Meyerriecks is less worried about rival nation states might use AI to outflank the United States than about getting U.S. leaders to believe what AI is telling them. "If I want to increase [certainty in a particular AI-aided assessment] what goes into it? What do I need in order to make a really good assessment on the back-end because that tells me what sort of collection I need to raise confidence to go address national leadership?"

The CIA currently has 137 pilot projects directly related to artificial intelligence, Meyerriecks, the CIA's deputy director for science and technology, told the Intelligence and National Security Summit in downtown DC.

These "experiments" include everything from automatically tagging objects in video (so analysts can pay attention to what's important) to better predicting future events based on big data and correlational evidence.

"Can we back into correlations with cause and effect that will allow us to be more predictive with

Trump's Pick for NSA/ CyberCom Chief Wants to Enlist AI For Cyber Offense

A look at Lt. Gen. Paul Nakasone's public statements about artificial intelligence, offense, and defense.

By Patrick Tucker

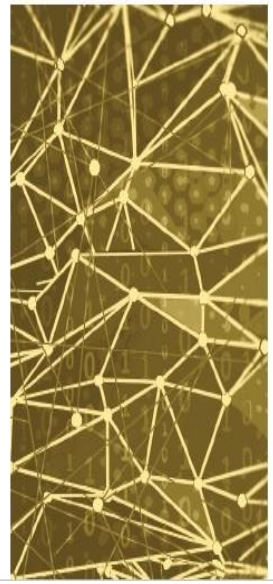
The Army general likely to be tapped to head U.S. Cyber Command and the NSA has some big plans for deploying cyber forces and using artificial intelligence in information attacks.

Lt. Gen. Paul Nakasone, who currently leads U.S. Army Cyber Command, is expected to nominated in the next few months to replace Adm. Michael Rogers, as first reported by The Cipher Brief (and confirmed by the *Washington Post* and a Pentagon source of our own). (**Update**: On January 30, Nakasone's nomination was further confirmed by *Politico*.)

How does Nakasone differ from his predecessor? The Army general's public comments suggest he would use artificial intelligence aggressively in offensive cyber operations, (read hacking activities.) That could mean a lot more offensive cyber activity to achieve an overwhelming effect and that, in turn, is a big shift from where the Defense Department was in 2016 when it treated public talk of cyber offensive operations with great delicacy.

Nakasone has pointed to the 2016 DARPA Grand Cyber Challenge as emblematic of how artificial intelligence will change both offensive and defensive operations, replacing human hackers with software that can heal its own bugs and vulnerabilities while simultaneously searching for and exploiting bugs in adversary systems.

"It's really easy to say, 'I'm going to get on this network, achieve presence, have a persistent ability to go after whatever I do in the future.' But what if you had a machine that did that? That was able to rapidly find a



MF3d, istock.com

China Is On a Whole-of-Nation Push for AI. The US Must Match It

Beijing is harnessing government and commercial entities in pursuit of a once-in-a-generation technological kingmaker.

By Elsa B. Kania

China has made no secret of its ambitions to lead the world in artificial intelligence, nor of the military and geopolitical advantage it hopes to gain from this rapidly advancing technology. A closer look at Beijing's whole-of-nation AI strategy shows the challenge to the United States — and suggests what America must do lest it be eclipsed in this latest round of great-power competition.

China's vision came into focus over the summer with the release of the New Generation AI Development Plan, which articulates an ambitious agenda to "lead the world" in the field. Chinese leaders, no longer content to copy Western technologies, are aiming to become the world's "premier AI innovation center," advancing an "innovation-driven" strategy for civilian and military development.

The implementation of this agenda will be a whole-of-government endeavor involving 15 central agencies and a growing number of local governments. Their efforts will foster the growth of a robust AI industry and ecosystem and pour billions into longer-term research and development of next-generation technologies. The plan will tap the dynamism of national tech champions, such as Baidu, Alibaba, Tencent, and iFlytek, that have been leading China's AI revolution.

Under the national strategy of "military-civil fusion," their breakthroughs can also be put to military use.

It is telling that the agencies responsible for the plan include the Central Military-Civil Fusion Development Commission and both the Equipment Development Department and Science and Technology Commission of the Central Military Commission, or CMC. The strategic advisory committee responsible for supporting the new plan's implementation also includes several PLA civilians (in uniform). And China's AI agenda has the very highest support, in the country's leader — and CMC Chairman — Xi

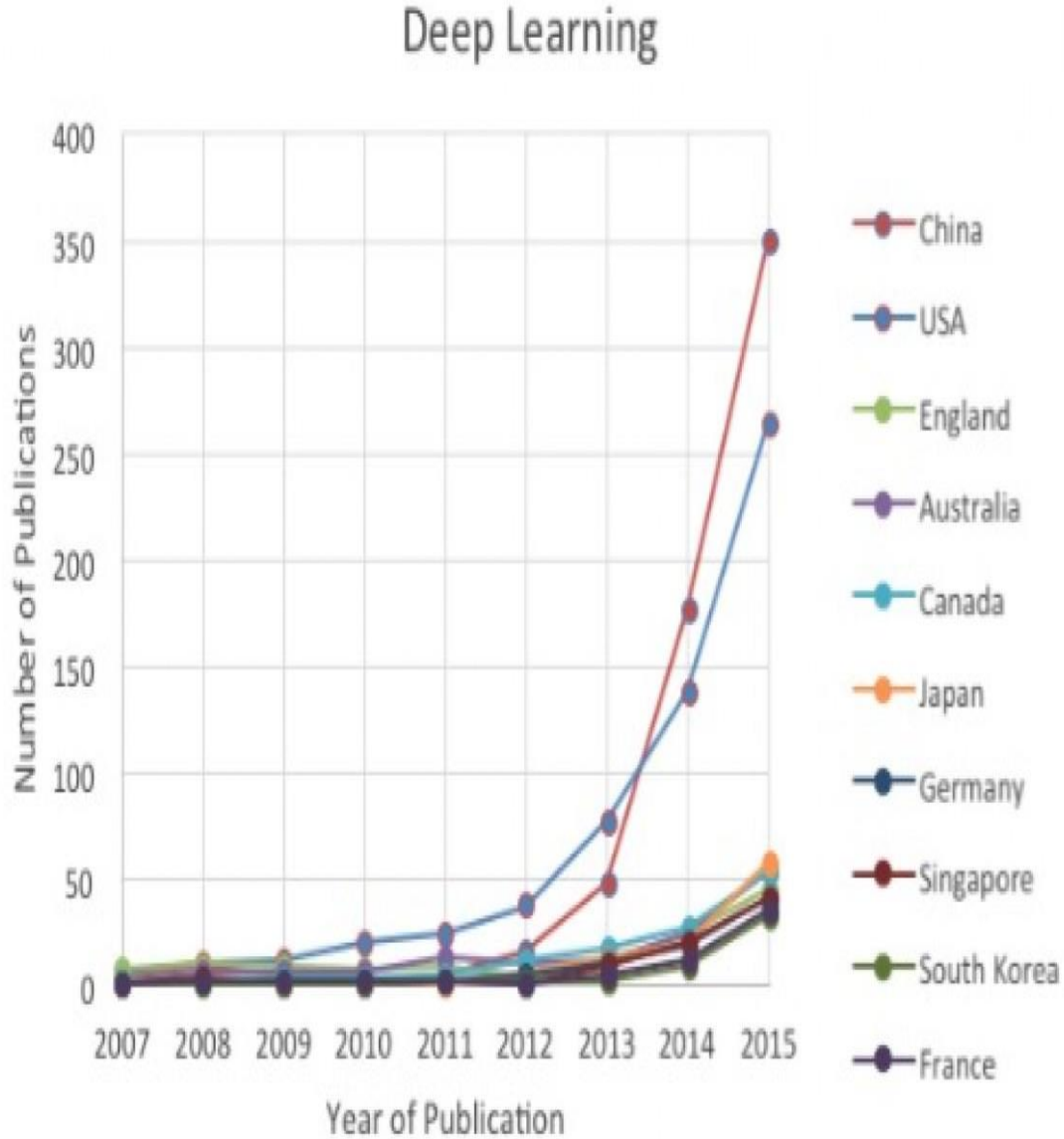


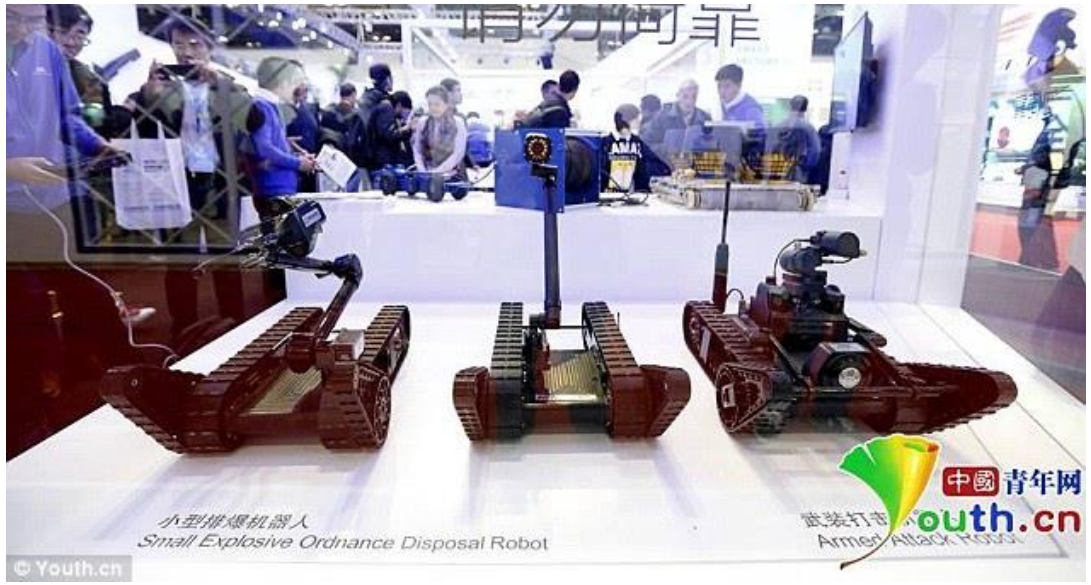
Figure 1: Journal articles mentioning "deep learning" or "deep neural network", by nation.[62]

Local governments charge ahead with AI development

Provincial and municipal plans set ambitious measures and goals

* targets specified for 2022

Sources: National, provincial, and municipal-level government documents


小型排爆机器人
Small Explosive Ordnance Disposal Robot
© Youth.cn
中國青年网
outh.cn


深瞳人眼摄像机动态人脸识别系统
WSJ
2017-02-27 下午 12:45:08 星期一
深瞳人眼摄像机
全景视频
人脸抓拍
抓拍图
特征库
相似度 86%
个人信息
李超
相似度 85%
个人信息
乐爱兵
相似度 86%
个人信息
王一舒
相似度 91%
04:38 / 08:10


NOBODY

China's brightest children are being recruited to develop AI 'killer bots'

- Beijing Institute of Technology recruits 31 'patriotic' youngsters for new AI weapons development programme
- Expert in international science policy describes course as 'extremely powerful and troubling'

« Le leader en intelligence artificielle dominera le monde »
Discours de Vladimir Poutine du 01 septembre 2017

Russian President Vladimir Putin, center, applauds during a meeting with students in Yaroslavl, Russia, Friday, Sept. 1, 2017. Putin said that whoever reaches a breakthrough in developing artificial intelligence will come to dominate the world. Putin, speaking Friday at a meeting with students, said the development of AI raises "colossal opportunities and threats that are difficult to predict now." (Alexei Druzhinin, Kremlin Pool Photo via AP) (Associated Press)

Russia Is Poised to Surprise the US in Battlefield Robotics

How? It's a story of leaders' unusual agreement, a focus on fast-and-cheap production, and a decision to field lethal robots for combat.

By Samuel Bendett

No one would call Russia's government and budgetary bureaucracy particularly nimble, nor its defense industry particularly advanced. Certainly, it trails Western economies in such key areas as communication equipment, microelectronics, high-tech control systems, and other key technologies. But in certain aspects of the field of unmanned military systems, Russia may be inching ahead of its competition in designing and testing a wide variety of systems and conceptualizing their future use.

In recent years, an unusually close alignment of its executive leadership and the Ministry of Defense on the importance of unmanned systems has vastly streamlined their funding, development, and deployment. (The defense minister has a direct line to the president, and final military decisions are often made by a very small circle of individuals—a far cry from the American budgetary process. As well, the Russian defense budget will remain largely unchanged over the next few years, give or take a few percent, even as other government ministries fight for budget share.)

Russia's swift progress in unmanned systems suggest that the United States and its allies should prepare for battle against foes who can put U.S. forces at a disadvantage by inhibiting their operational capabilities.

Air, Land, and Sea

The vast military force that Russia inherited from the Soviet Union was generally older and less technologically advanced than the U.S. arsenal, but it did include a relatively good ISR UAV: the Pchela/Shmel, which has been used in every major conflict from the Chechen wars in the 1990s to today's Syrian campaign. In the 2000s, Russia compensated for its lack of domestic UAV manufacturing capability by importing Forpost and Zastava UAVs from Israel. Today, the Eleron, Orlan, and Forpost trio of UAVs are in widespread use by Russian forces, including

V - Stratégie et Initiatives françaises en IA

France IA 1 - Mission Villani – Hub France IA – programme SécurIA

Groupe Sécurité-IA

Du Think Tank

Au

Do Tank

Le groupe « Sécurité IA » du Hub

Co-animé par Eric Hazane & Thierry Berthier

Un noyau dur d'une vingtaine de membres actifs pour le moment incluant Air Liquide, SNCF, EDF, MinInt, DGA, ANSSI, EMM, des startups et ETI , ITRUST, SNIPS, NXU, des chefs d'entreprises, un Hub IA Toulousain, Pôle d'Excellence Cyber Bretagne (PEC) & IMT Atlantique.

Une veille Sécurité – IA diffusée deux fois par semaine et mise en ligne sur un site wordpress :

https://iasecurite.wordpress.com/

Veille extraite de la veille cyber active depuis 6 ans :

https://veillecyberland.wordpress.com/

Hub France IA :

http://www.hub-franceia.fr/

Groupe Sécurité-IA

Le programme SécurIA

AMI Mutualisation de l'IA - Bercy