Cyber security

from technological research to offensive (mis)use

Erik Zouave

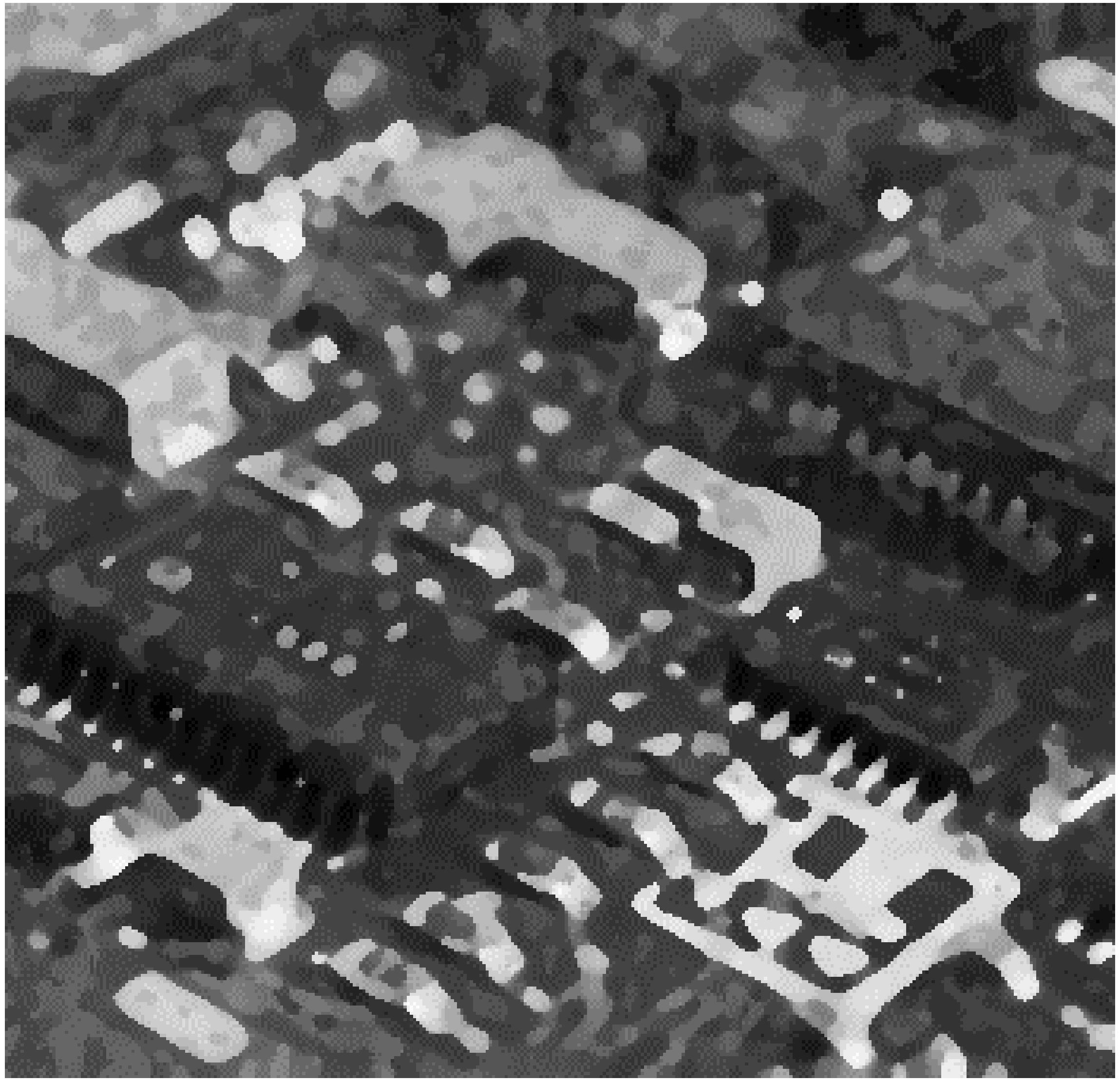
Disposition

1. Background

2. Definitions & Examples

3. Regulatory Approaches

4. General Mitigations



Background

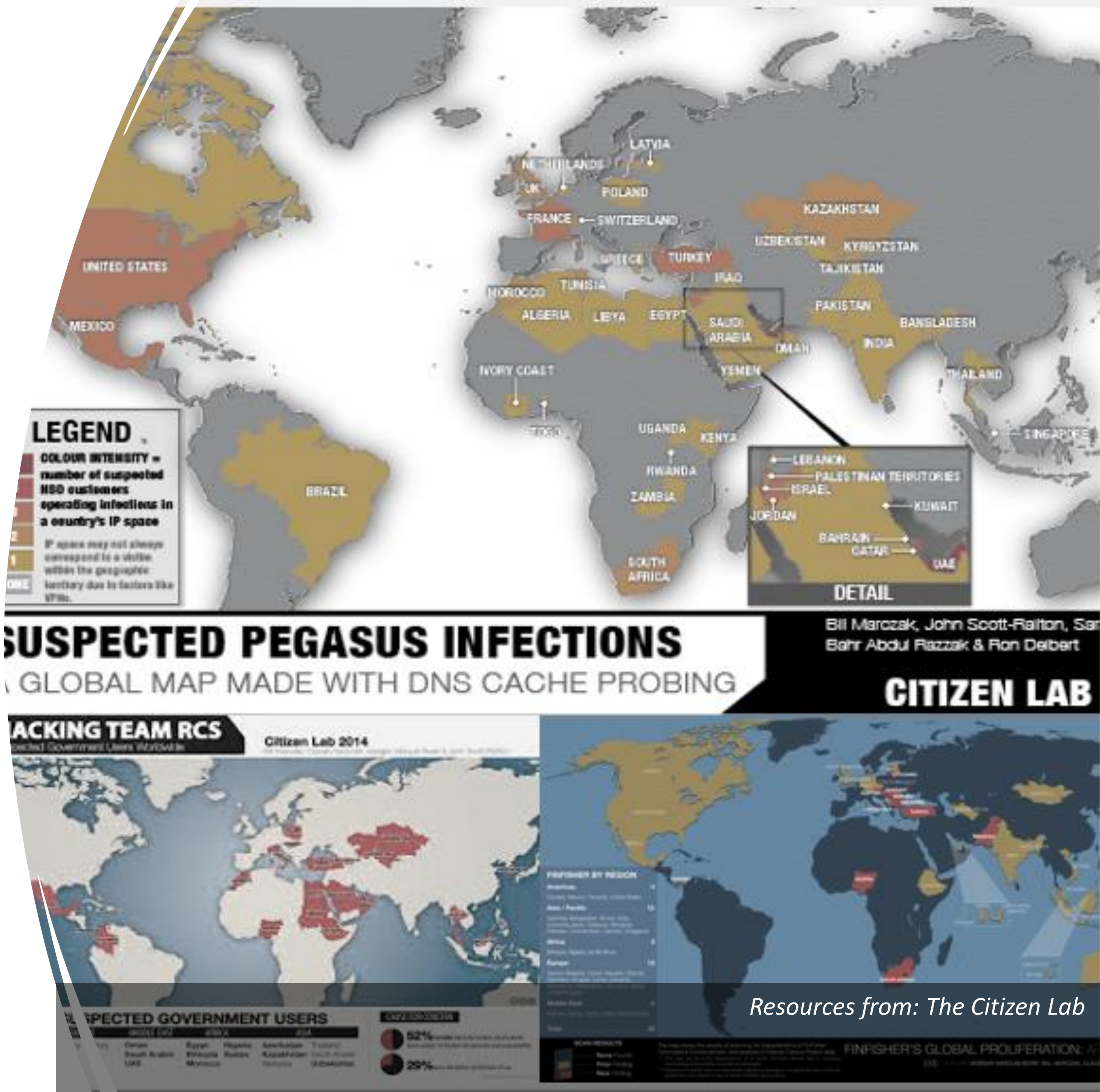
Flashback

- **Arab Spring:** surveillance targeting critics and dissidents.
- **Technology flow:** from western democracies to areas with repressive policies.
- **Initial focus:** human rights and export controls.
- **Primary sources:** technical investigations and unlawful leaks.

Marczak et al. (2014). “When Governments Hack Opponents: A Look at Actors and Technology”. In *Proceedings of the 23rd USENIX Security Symposium.* San Diego; Marquis-Boire, (2012). “Backdoors are Forever: Hacking Team and the Targeting of Dissent?”. *Research Brief October 2012*

Controversies & complaints

- ”Lawful Hacking” is especially marred by complaints
- 8 complaints against intermediaries
- 2 complaints against customers (governments)
- Spanning 8 years
- Stretching across 5 jurisdictions
- Regarding export control compliance, harassment, varying forms of cybercrime



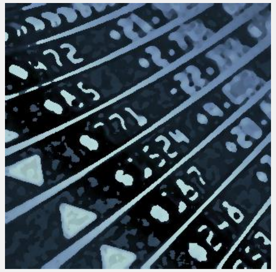
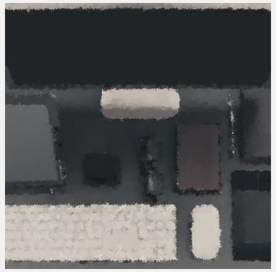
LEGEND
COLOUR INTENSITY = number of suspected NSO customers operating infections in a country's IP space
IP space may not always correspond to a victim within the geographic territory due to factors like VPNs.
NONE
UNITED STATES
MEXICO
BRAZIL
NETHERLANDS
LATVIA
UK
POLAND
FRANCE
SWITZERLAND
GREECE
TURKEY
KAZAKHSTAN
UZBEKISTAN
KYRGYZSTAN
TAJIKISTAN
MOROCCO
TUNISIA
IRAQ
ALGERIA
LIBYA
EGYPT
SAUDI ARABIA
OMAN
PAKISTAN
INDIA
BANGLADESH
IVORY COAST
TOGO
YEMEN
THAILAND
UGANDA
KENYA
RWANDA
ZAMBIA
SOUTH AFRICA
SINGAPORE
LEBANON
PALESTINIAN TERRITORIES
ISRAEL
JORDAN
KUWAIT
BAHRAIN
QATAR
UAE
DETAIL
SUSPECTED PEGASUS INFECTIONS
A GLOBAL MAP MADE WITH DNS CACHE PROBING
Bill Marczak, John Scott-Railton, Sar
Bahr Abdul Razzak & Ron Deibert
CITIZEN LAB
HACKING TEAM RCS
Suspected Government Users Worldwide
Citizen Lab 2014
SUSPECTED GOVERNMENT USERS
52%
29%
FINFISHER BY REGION
FINFISHER'S GLOBAL PROLIFERATION

Resources from: The Citizen Lab

What does this have to do with researchers?

- Insights into security, insecurity and mitigations.
- Developers of technologies.
- Drivers in technology maturity.
- Transparent, relatively unprotected and at risk of having their work exploited.
- In the employ of security industry and authorities.
- Potential knowing or unknowing collaborators with repressive interests.
- Researchers can also commit crimes with technologies.

Examples of where can misuse start?

Researcher icon	✓		✓	✓	✓	✓
Examples of Actors						
	Individuals	Criminal organizations	Brokers	Research institutes	Industry	Government
Examples of “bad actors”	✓ Careless. ✓ Uninformed. ✓ Criminal.	✓ Criminal. ✓ State-sponsored. ✓ Terrorists.	✓ Lack of due diligence. ✓ Offense-accepting. ✓ Embargo-busting.	✓ Repressive funders.	✓ Lack of due diligence. ✓ Embargo-busting. ✓ Repressive ownership.	✓ Repressive regimes. ✓ Allies & supporters of repressive regimes.
Examples of dissemination	✓ Careless publication. ✓ Irresponsible disclosure. ✓ Brokers. ✓ Darknet.	✓ Brokers. ✓ Industry. ✓ Government. ✓ Darknet.	✓ Government. ✓ Industry.	✓ Careless publication. ✓ Insecure handling. ✓ Funders. ✓ Consortia.	✓ Insecure handling. ✓ Customers.	✓ “Allies”.

Definitions & Examples

Offensive cyber technologies

Characteristics of terminology

Type of definition		- Academic - Doctrinal (military) - Lexical (crime)
Examples of definitions	***Military***	Arguably based on **context, timing, intent** and **behavio**r in the case of government activities. E.g., operations to: - **project power** (U.S.; U.K.) - **gain momentum** and take initiative to **attain interest** or - **achieve goals** (U.K.; SWE), - **influence** or **preempt** actions (NL). Involves **feints** and **exploitations** (US).
	Crime	(Intending to) **attack**(ing) someone with a **weapon**.
Relation to misuse		May vary: - Competition between adversaries (military) - Criminal misuse

Technical characteristics

Deny

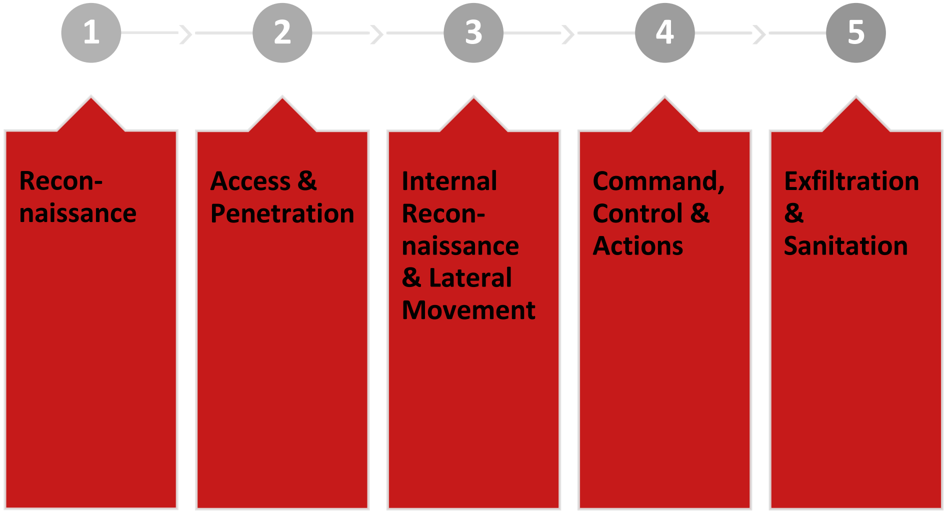
Disrupt

Degrade

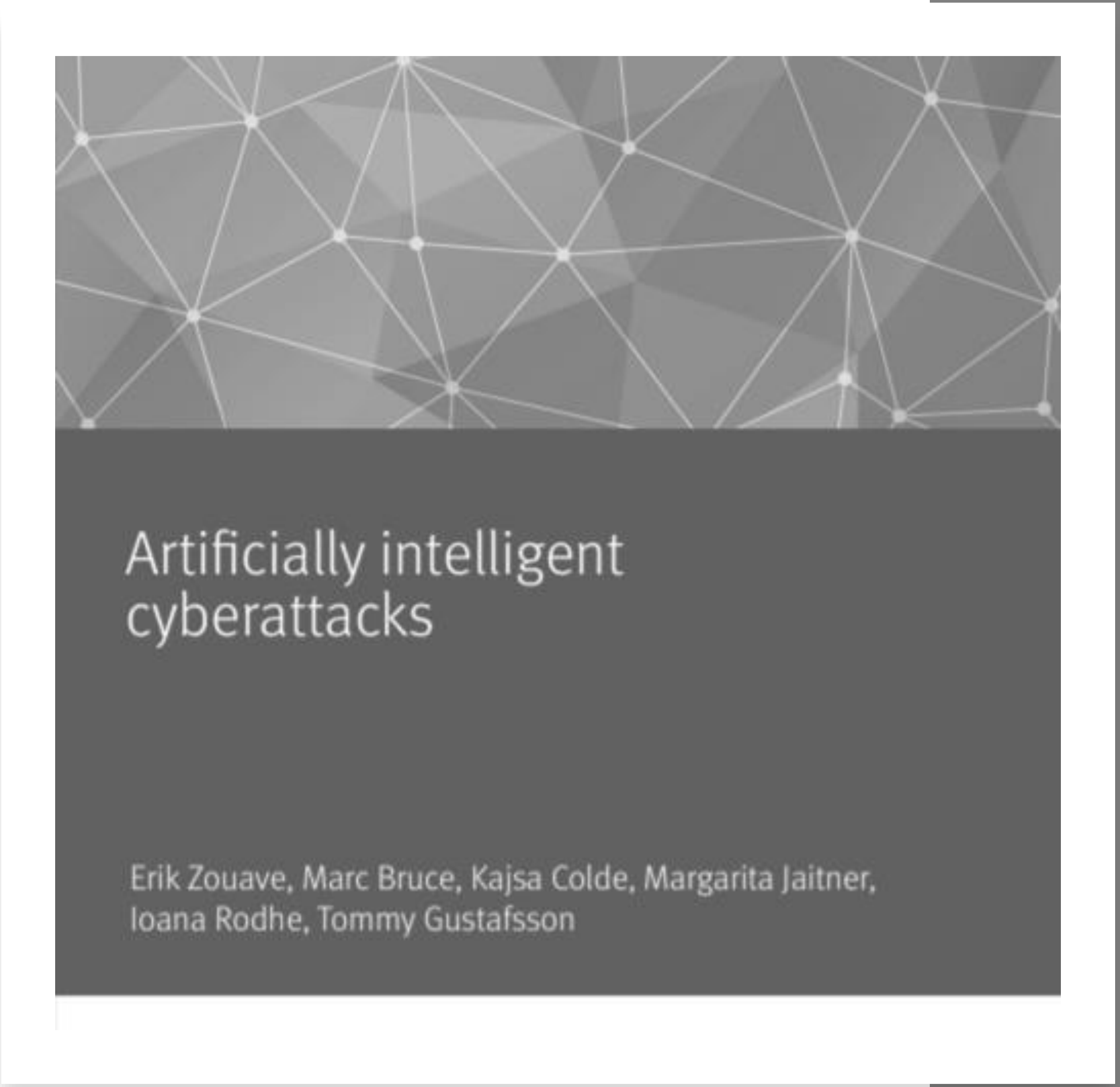
Destroy

Uren, Hogeveen, Hansson. (2018). Defining offensive cyber capabilities;
Mladenovic, Radunovic. (2018). Defining offensive cyber capabilities;
Lin, Zegart. (2018). Bytes, Bombs and Spies;
Lexico; Cambridge Dictionary.

Kill chains, anatomies, phases are familiar



1. Reconnaissance
2. Access & Penetration
3. Internal Reconnaissance & Lateral Movement
4. Command, Control & Actions
5. Exfiltration & Sanitation

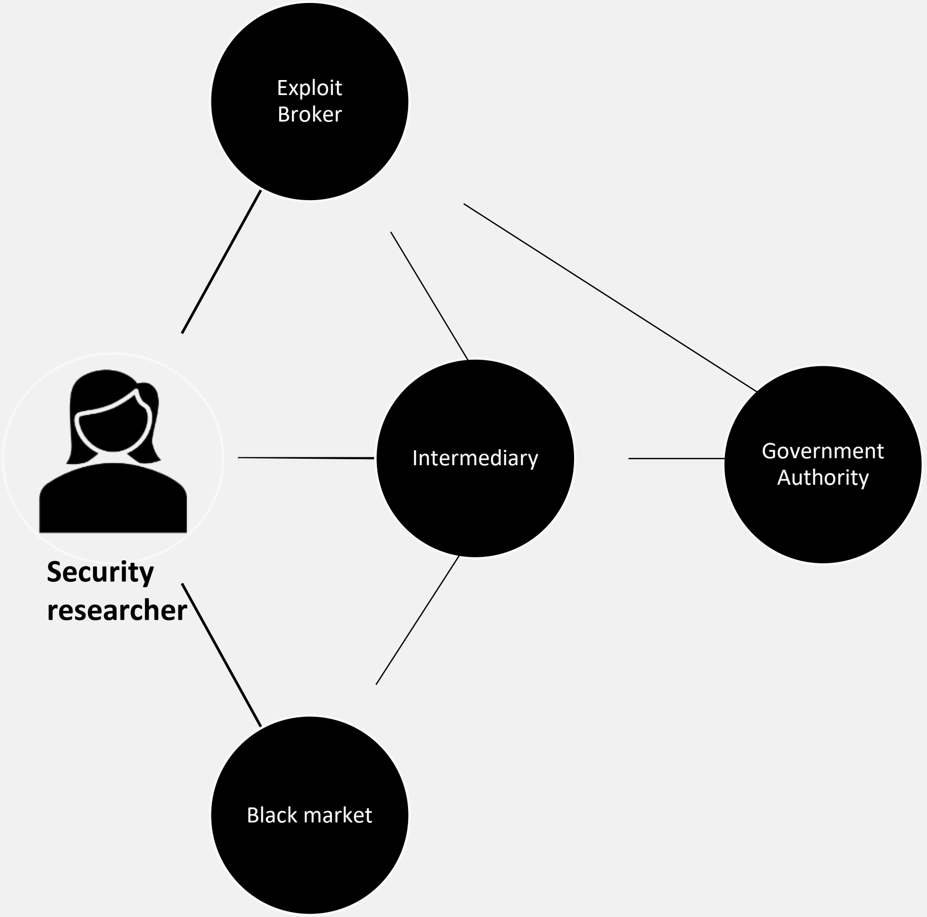


Artificially intelligent cyberattacks

Erik Zouave, Marc Bruce, Kajsa Colde, Margarita Jaitner, Ioana Rodhe, Tommy Gustafsson

Zouave et al. (2020). Artificially intelligent cyberattacks.

Examples: offensive cyber industry – from vulnerability to offensive tool



Broker/Market

- **Vulnerabilities**
- **Exploits**
- **0days**

Intermediary

- **Propagation methods:** e.g., man-in-the-middle, false wifi, Bluetooth, password cracking…
- **Payloads & remote commands:** e.g., data modification, real-time surveillance (keystrokes, audio, video, text)…
- **Monitoring systems:** e.g. text, language, and image analysis, object recognition, biometrics..
- **Training & consultancy**

Government

- **Offensive operations**

Vupen; Zerodium; Coseinc; Exodus; Netragard; Elaman; NSO Group; Cyberbit, Hacking Team; Gamma.

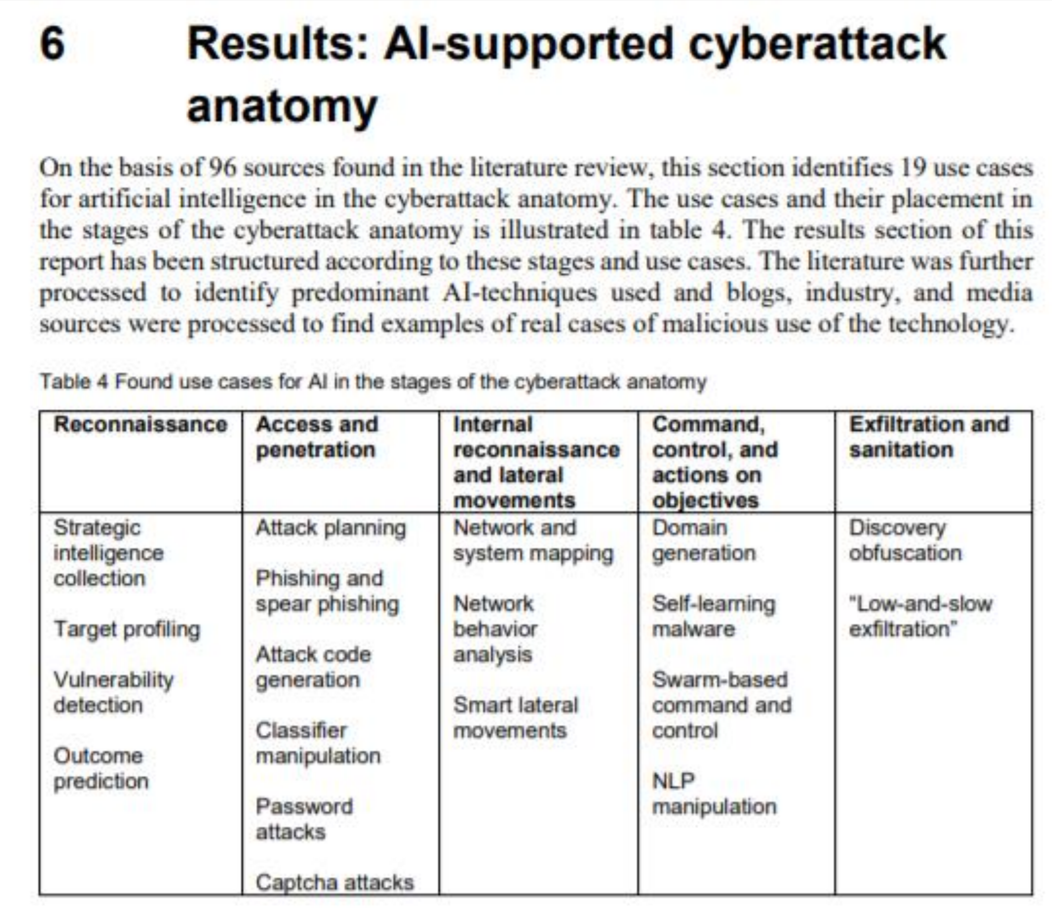
Examples: AI-supported offensive cyber

”CyberLover” (2007)

- Natural Language Processing (NLP)
- Profiled Russian dating chatroom users; e.g. “romantic lover”
- Adapted tailored dialogue options to profiles
- Provides fraudulent links
- Data theft
- Abt. 1 new relationship/ 3 min.

Zouave et al. (2020). Artificially intelligent cyberattacks;
Rossi. (2007). Beware the CyberLover that Steals personal Data.

AI-supported cyber offense overview

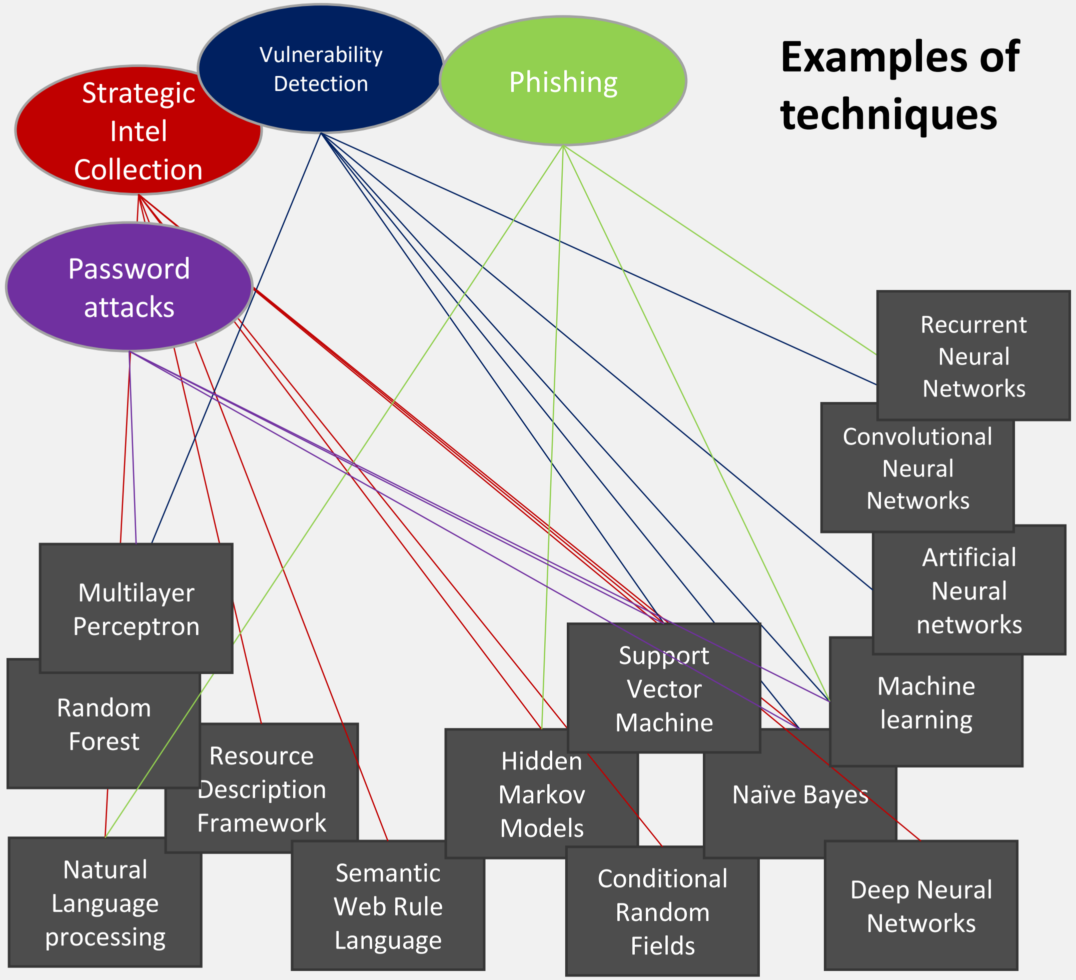


6 Results: AI-supported cyberattack anatomy

On the basis of 96 sources found in the literature review, this section identifies 19 use cases for artificial intelligence in the cyberattack anatomy. The use cases and their placement in the stages of the cyberattack anatomy is illustrated in table 4. The results section of this report has been structured according to these stages and use cases. The literature was further processed to identify predominant AI-techniques used and blogs, industry, and media sources were processed to find examples of real cases of malicious use of the technology.

Table 4 Found use cases for AI in the stages of the cyberattack anatomy

Reconnaissance	Access and penetration	Internal reconnaissance and lateral movements	Command, control, and actions on objectives	Exfiltration and sanitation
Strategic intelligence collection Target profiling Vulnerability detection Outcome prediction	Attack planning Phishing and spear phishing Attack code generation Classifier manipulation Password attacks Captcha attacks	Network and system mapping Network behavior analysis Smart lateral movements	Domain generation Self-learning malware Swarm-based command and control NLP manipulation	Discovery obfuscation "Low-and-slow exfiltration"

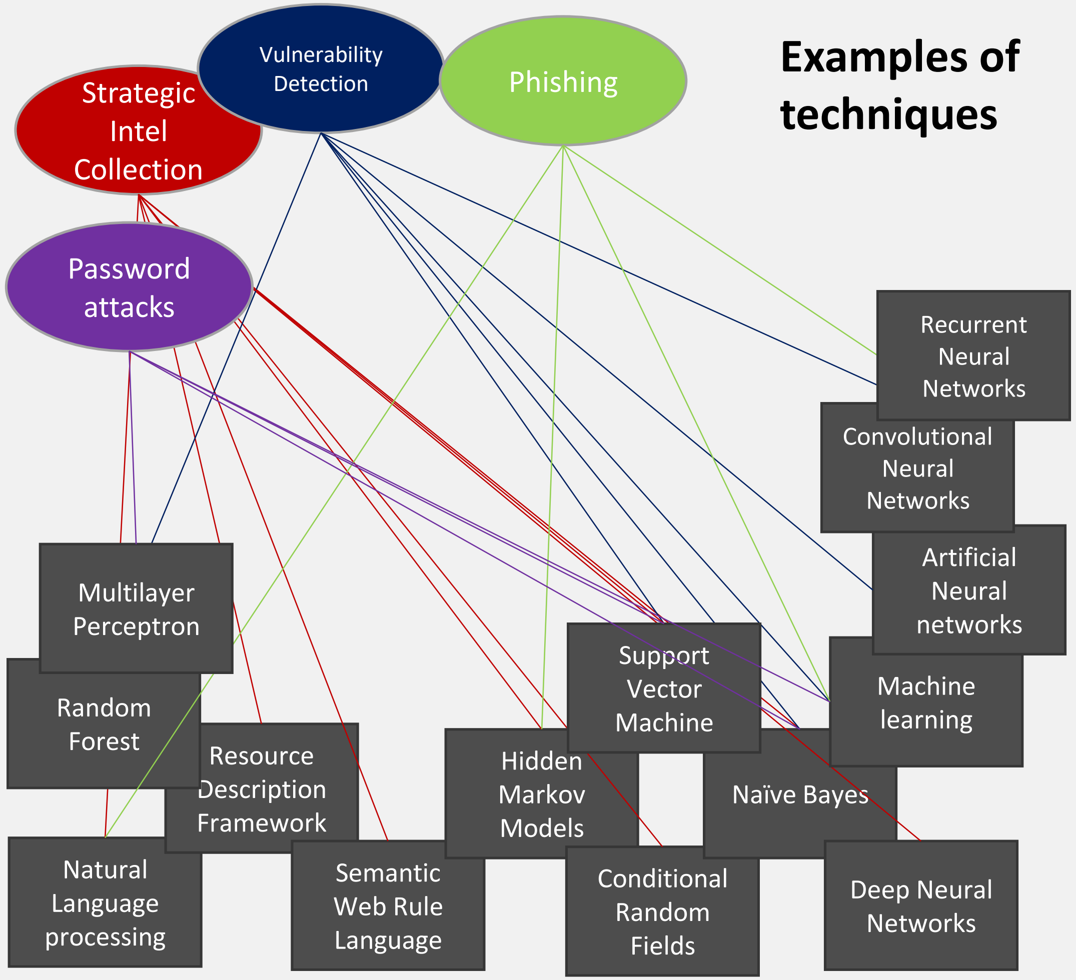

Strategic Intel Collection
Vulnerability Detection
Phishing
Password attacks
Recurrent Neural Networks
Convolutional Neural Networks
Artificial Neural networks
Multilayer Perceptron
Random Forest
Resource Description Framework
Support Vector Machine
Hidden Markov Models
Machine learning
Naïve Bayes
Natural Language processing
Semantic Web Rule Language
Conditional Random Fields
Deep Neural Networks

Examples of techniques

Zouave et al. (2020). *Artificially intelligent cyberattacks.*

AI-supported cyber offense overview

- Where do you draw the line for offensive means and methods?
- How do you identify the risk of misuse?
- Which technologies, techniques, methods, know-how?


Examples of techniques
Strategic Intel Collection
Vulnerability Detection
Phishing
Password attacks
Recurrent Neural Networks
Convolutional Neural Networks
Artificial Neural networks
Machine learning
Multilayer Perceptron
Random Forest
Resource Description Framework
Natural Language processing
Semantic Web Rule Language
Hidden Markov Models
Support Vector Machine
Naïve Bayes
Conditional Random Fields
Deep Neural Networks

Zouave et al. (2020). Artificially intelligent cyberattacks.

Zouave et al. (2020). Artificially intelligent cyberattacks.

Legislative Approaches

”Misuse” in the sense of export controls

Considering the emergence of new categories of dual-use items, and in response to calls from the European Parliament and indications that ***certain cyber-surveillance technologies exported*****[, transfer, brokering or transit]** ***from the Union*** *have been* ***misused*** *by persons complicit in or responsible for directing or committing* ***serious violations of human rights or international humanitarian law*** *in situations of armed conflict or* ***internal repression,*** *it is appropriate to control the export of those technologies in order to protect public security as well as public morals.*

(Rec 5, COM (2016) 616)

§

Approach in Export Controls

Regulation 428/2009

Intrusion Software

Regulated end-use

- ✓ Military (art 4),
- ✓ Embargoed territory (art 4),
- ✓ Public security threat (art 8),
- ✓ Human rights abuse (art 8).

Items/info

- ✓ Equipment,
- ✓ Technology, and
- ✓ Software
- ✓ Know-how

Capabilities

- ✓ Avoids detection, or
- ✓ Defeats portion, and
- ✓ Extracts data/information,
- ✓ Modifies system/user data, or standard execution path of a program or process, to
- ✓ Allow the execution of externally provided instructions.

Additions

COM (2016) 616

Cyber-Surveillance Tools

Status:

- ✓ Council and Parliament provisional agreement
- ❑ COREPER endorsement
- ❑ Readings
- ❑ Conciliation
- ❑ Result

Regulated end-use

- ✓ Serious human rights violations (art 4),
- ✓ Serious violations of humanitarian law (art 4),
- ✓ Threat to international security (art 4),
- ✓ Threat to essential security interests of EU and MS (art 4),
- ✓ Terrorism (art 4),

Capabilities

- ✓ Enable the covert intrusion into information and telecommunication systems.
- ✓ Monitor, extracts, collects, or analyses data.
- ✓ Incapacitates or damages the system.

Challenges to Export Controls

Challenges

- Limited to the like-minded
- Subject to evasion
- Ineffective national implementation and oversight

Implications for research

1. Only partner and collaborate with or disseminate to the like-minded.
2. Adopt additional controls when collaborating with the like-minded but ineffective.
3. Assess the reliability of partners on a case-to-case basis.

”Misuse” in the sense of cybercrime

Each Party shall adopt such legislative and other measures as may be necessary to establish as criminal offences under its domestic law, when committed ***intentionally and without right****:*

a the ***production, sale, procurement*** *for use,* ***import, distribution or otherwise making available of****:*

> *i a device, including a computer program, designed or adapted* ***primarily for the purpose of committing any of the offences*** *established in accordance with Articles 2 through 5;*
>
> *ii a computer password, access code, or similar data by which the whole or any part of a computer system is capable of being accessed,*
>
> *with intent that it be used for* ***the purpose of committing any of the offences*** *established in Articles 2 through 5; and*

b the possession of an item referred to in paragraphs a.i or ii above, with intent that it be used for the purpose of committing any of the offences established in Articles 2 through 5. A Party may require by law that a number of such items be possessed before criminal liability attaches.

(art 6 on “Misuse of devices”, Convention on Cybercrime ETS. 185)

§

Approach in Cybercrime Law

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

Convention on Cybercrime ETS 185
Directive 2013/14/EU

Cybercrime

Regulated end-use	Items/info	Capabilities
✓ Illegal access to information systems (art 2/3), ✓ Illegal system interference (art 5/4), ✓ Illegal data interference (art 4/5).	✓ Device (art 6, ETS185), ✓ Computer program (art 6, ETS185), ✓ Passwords, access codes or similar data (art 6, ETS185), ✓ Tools (art 7, Dir2013/14/EU), ✓ Incitement, aiding and abetting (art 8, Dir2013/14/EU).	✓ Access without right, ✓ Seriously hindering or interrupting the functioning of an information system by inputting computer data, by transmitting, damaging, deleting, deteriorating, altering or suppressing such data, or by rendering such data inaccessible, intentionally and without right, ✓ Deleting, damaging, deteriorating, altering or suppressing computer data on an information system, or rendering such data inaccessible, intentionally and without right

Challenges to Cybercrime Law

Challenges

- Does not apply to public bodies (state ”hacking”)
- Does not apply where there is authorization under national law
- Mere status as a public body or authorization is no guarantee against human rights abuse or security threats

Implications for research

1

Assess end-use, e.g. ECHR art 8 ”necessary in a democratic society”.

2

Assess partner’s associations, e.g.

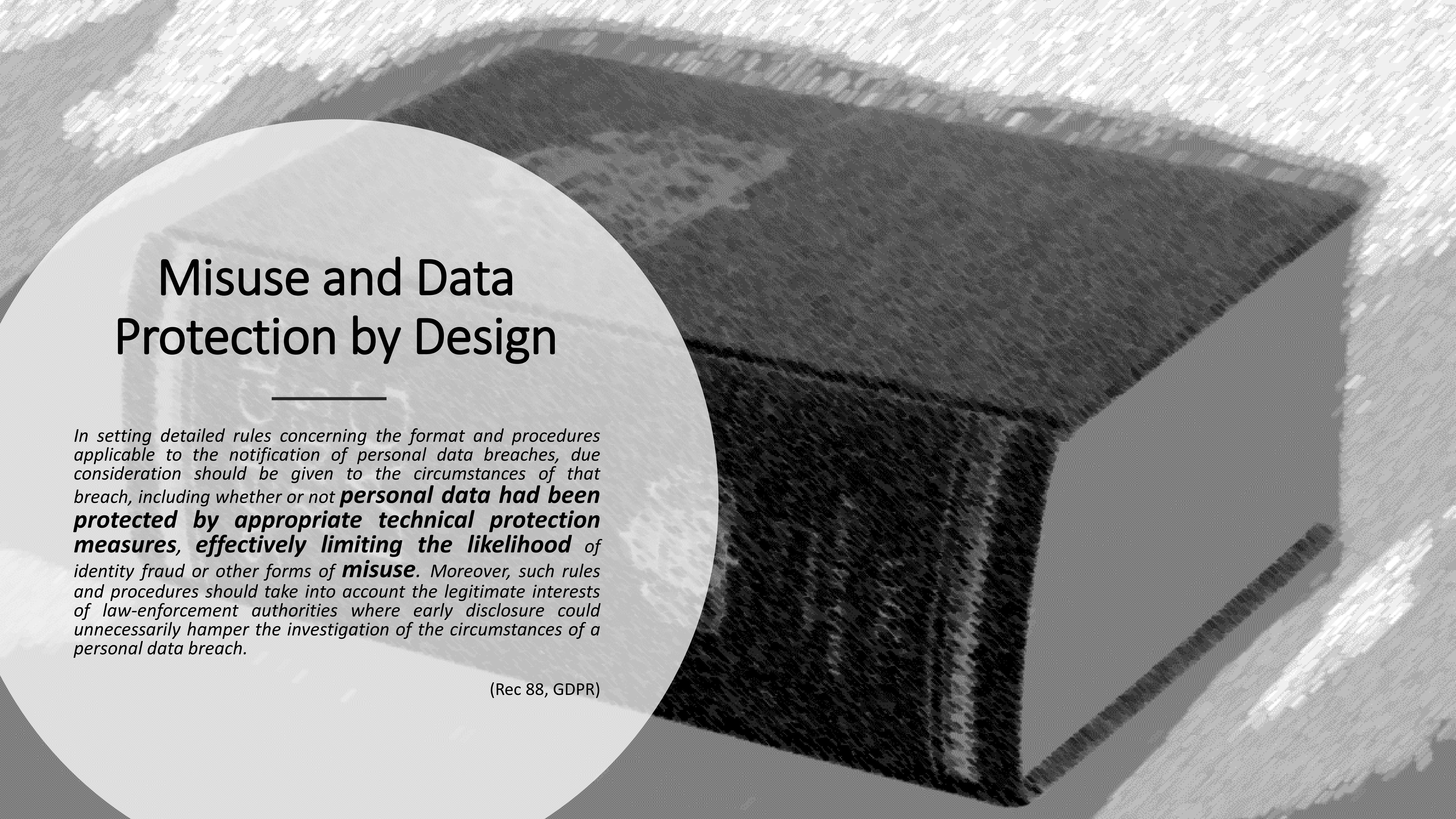
- EDPB “essential guarantees”,

- Known use of cyber tools against EU targets etc.

3

Personnel checks and other controls on staff.

Blinderman, Din, (2017). “Hidden by Sovereign Shadows: Improving the Domestic Framework for Deterring State-Sponsored Cybercrime”. *Vanderbilt Journal of Transnational Law* 50:1



Misuse and Data Protection by Design

In setting detailed rules concerning the format and procedures applicable to the notification of personal data breaches, due consideration should be given to the circumstances of that breach, including whether or not ***personal data had been protected by appropriate technical protection measures****,* ***effectively limiting the likelihood*** *of identity fraud or other forms of* ***misuse****. Moreover, such rules and procedures should take into account the legitimate interests of law-enforcement authorities where early disclosure could unnecessarily hamper the investigation of the circumstances of a personal data breach.*

(Rec 88, GDPR)

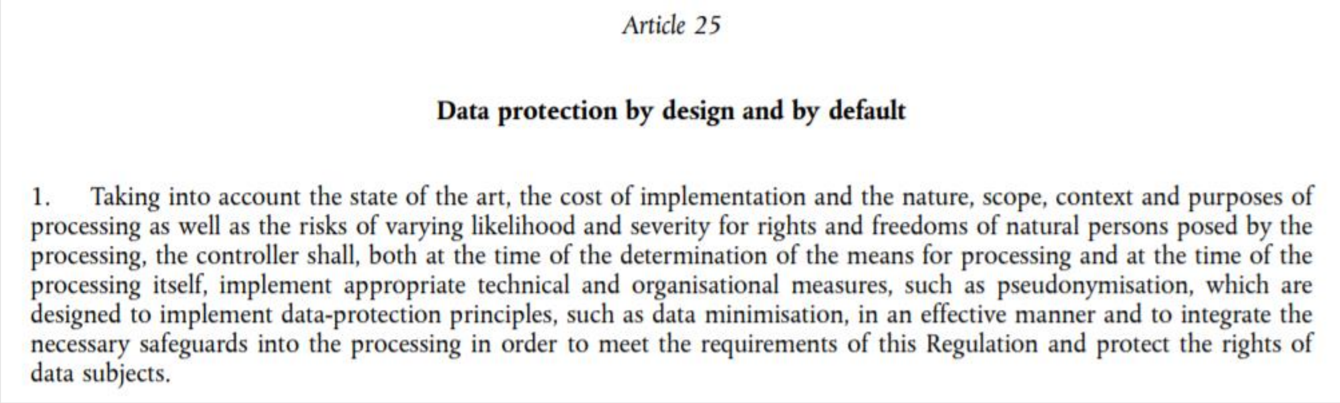
§

By Design Approaches

Regulation EU 2016/679
Directive EU 2016/680

Data Protection by Design

Regulated end-use	Items/info	Capabilities
✓ Processing with risks, e.g. to rights and freedoms.	✓ Personal data ✓ Processing systems	Processing of personal data



Article 25

Data protection by design and by default

1. Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.

Challenges to Data Protection by Design

Challenges

- Limits on the scope of data protection, e.g. national security
- European Data Protection Supervisor’s calls for discussions were ignored by EU law enforcement
- Processors are frequently not designers

Implications for research

1

Apply the “by design” criterion on national security projects anyway, and

2

Assess the risk of data processing innovations being adopted without integration of “by design” considerations down the line, or

3

Avoid association with the combination of offensive technologies, know-how and national security altogether.

Bygrave. (2017). Data Protection by Design and by Default: Deciphering the EU’s Legislative Requirements. Oslo Law Review 4:2;
EDPS. (2015). Opinion 8/2015 Dissemination and use of intrusive surveillance technologies.

BTW: ”By Design” is also for Humanitarian Law & Human Rights in Armed Conflict

In the ***study, development, acquisition or adoption of a new weapon, means or method of warfare****, a High Contracting Party is under an obligation to* ***determine whether its employment*** *would, in some or all circumstances, be* ***prohibited by this Protocol or by any other rule of international law*** *applicable to the High Contracting Party.*

(Article 36 on “New Weapons”, Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977)

Other Possible Legislative Approaches

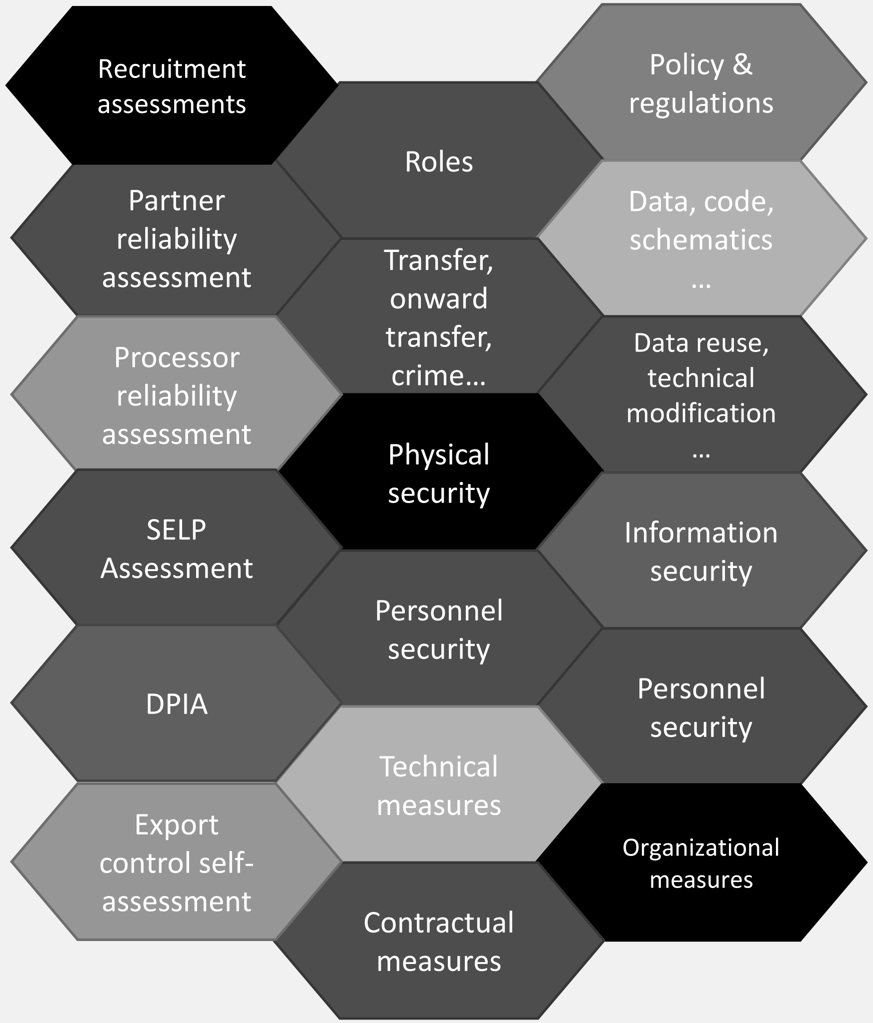
- ✓ Intelligence regulations
- ✓ Counter-espionage regulations
- ✓ Vulnerability equities regulations
- ✓ Public procurement regulations
- ✓ Investment screening and foreign takeovers regulations
- ✓ Confidentiality regulations
- ✓ IT-outsourcing regulations
- ✓ Research ethics regulations

Many links on the legal chain…

General Mitigations

Apply Aspects of Information Security Management to Tech Research at Risk

- ✓Establish responsibility
- ✓Classify Information Assets
- ✓ Assess risk, severity and likelihood
- ✓ Adopt appropriate mitigations
- ✓ Report incidents to appropriate authorties
- ✓ Evaluate, revise and improve



Recruitment assessments

Partner reliability assessment

Processor reliability assessment

SELP Assessment

DPIA

Export control self-assessment

Roles

Transfer, onward transfer, crime…

Physical security

Personnel security

Technical measures

Contractual measures

Policy & regulations

Data, code, schematics …

Data reuse, technical modification …

Information security

Personnel security

Organizational measures

Protect the Gravensteen Defenses!

Impromptu Medieval Belgian Infosec Wargame

You are the Chief Strategist of Gravensteen's Defense. Identify:

1. **One "dual use" asset** (both defensive and offensive) to protect from misuse: *human, organizational, technical etc*.
2. **One risk** associated with your asset.
3. **One key factor to evaluate** the severity or likelihood of the risk.
4. **One appropriate mitigation** for your risk.

Be on the “good” side of tech – Merry Christmas!