Artificial Intelligence in Cybersecurity

Olga Tushkanova

Laboratory of Computer Security Problems, SPIIRAS

Saint Petersburg, 2019

Content

1. What is Artificial Intelligence (AI)?
2. Tasks and methods of AI.
3. Some statistics on AI.
4. Cybersecurity scope.
5. Cybersecurity lifecycle.
6. Red and blue AI in cybersecurity.
7. Blue AI in cybersecurity use cases.
8. AI Methods for Cybersecurity Use Cases.
9. Research of Laboratory of Computer Security Problems, SPIIRAS.

Artificial Intelligence is not easy to define

Intelligence is the ability to learn and solve problems.

[The Webster's Dictionary]

Artificial intelligence is the science and engineering of making intelligent machines.

[John McCarthy]

Artificial intelligence is the intelligence exhibited by machines or software.

[Wikipedia]

Artificial intelligence is the study and design of intelligence agents where an intelligent agent is a system that perceives it's environment and takes actions that maximizes its chances of success.

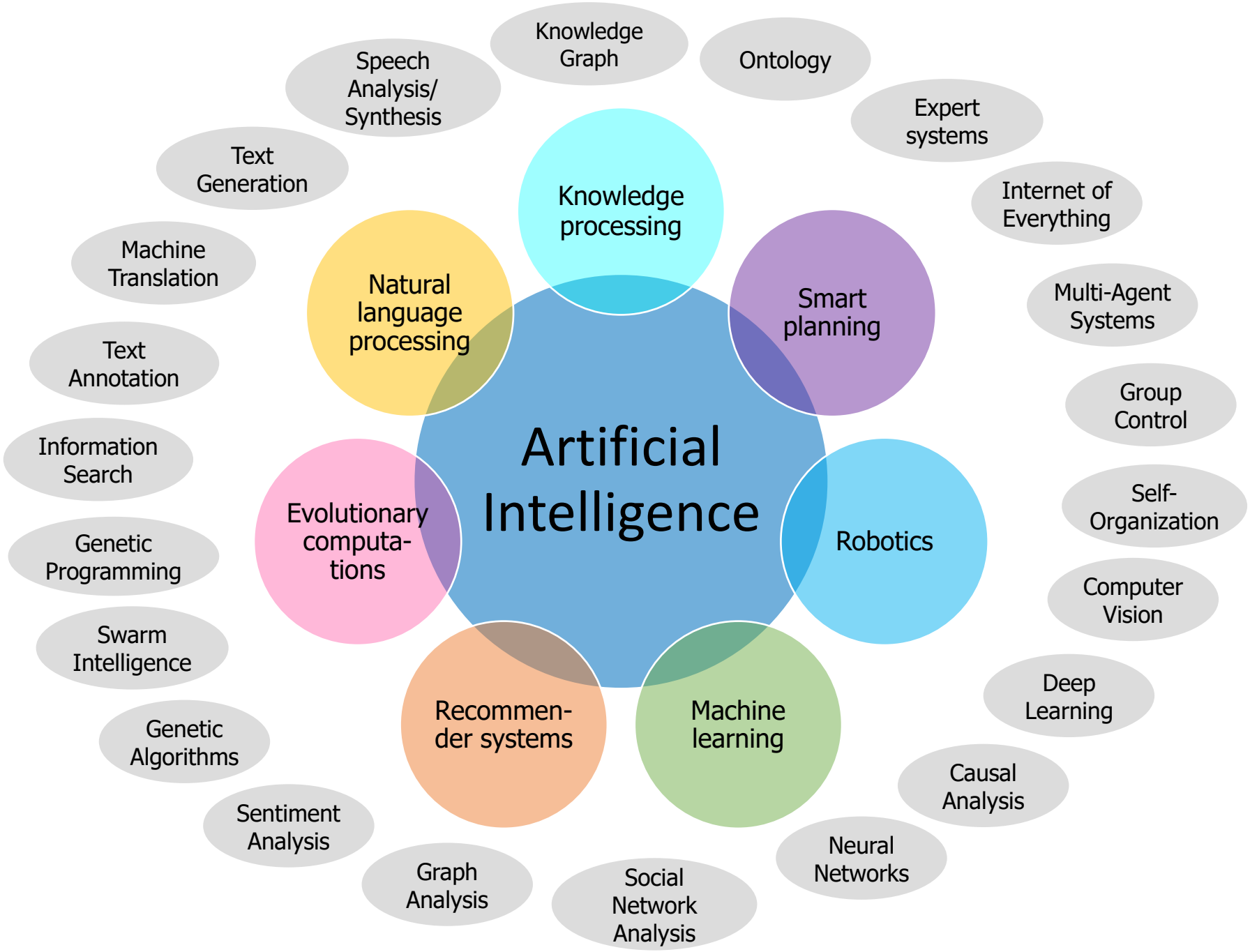
[Russell and Norvig]

	Human in the loop →	No human in the loop
Hardwired systems	**Assisted Intelligence** AI systems that assist humans in making decisions or taking actions. Hardwired systems that do not learn from their interactions.	**Automation** Automation of manual and cognitive tasks that are either routine or non-routine. This does not involve new ways of doing things.
Adaptive systems	**Assisting Intelligence** AI systems that augment human decision making and continuously learn from their interactions with humans and the environment.	**Autonomous Intelligence** AI systems that can adapt to different situations and can act autonomously without assistance.

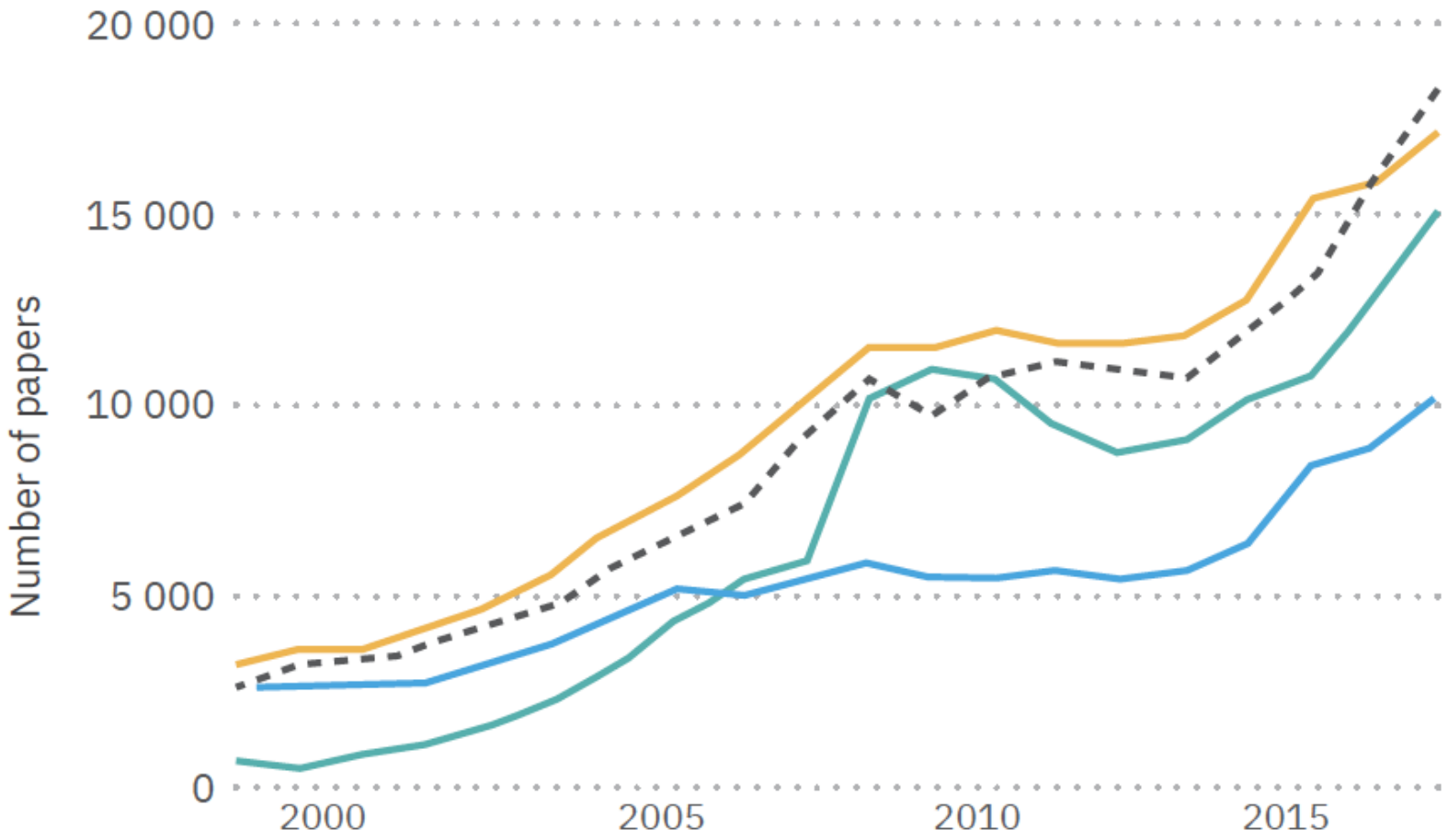
A set of technological and software solutions leading to a result similar to human intelligent activity and used to solve applied problems using computer vision systems, natural language processing, speech recognition and synthesis, recommender systems and intelligent decision support systems, as well as systems based on promising methods.

The Basics of Artificial Intelligence

Philosophy	Representation of knowledge, logic, mind as the basis of a physical learning system, rationality
Math	Algorithms, complexity of algorithms, convergence, probability theory, discrete mathematics, statistics, machine learning, graph theory
Economy	Expert systems, decision theory, game theory
Psychology	Behavioral analysis, experimental studies
Neuroscience and Brain Science	Learning, neural networks, augmented reality, neurobiology
Control Theory and Cybernetics	Information theory, robotics, building effective systems, knowledge representation and processing, ontology
Computer science	Operating systems, programming, databases management, networks
Linguistics	Grammar, vocabulary, natural language processing


Artificial Intelligence
Knowledge processing
Natural language processing
Smart planning
Evolutionary computa-tions
Robotics
Recommen-der systems
Machine learning
Knowledge Graph
Ontology
Expert systems
Internet of Everything
Multi-Agent Systems
Group Control
Self-Organization
Computer Vision
Deep Learning
Causal Analysis
Neural Networks
Social Network Analysis
Graph Analysis
Sentiment Analysis
Genetic Algorithms
Swarm Intelligence
Genetic Programming
Information Search
Text Annotation
Machine Translation
Text Generation
Speech Analysis/ Synthesis

Scopus AI Publications

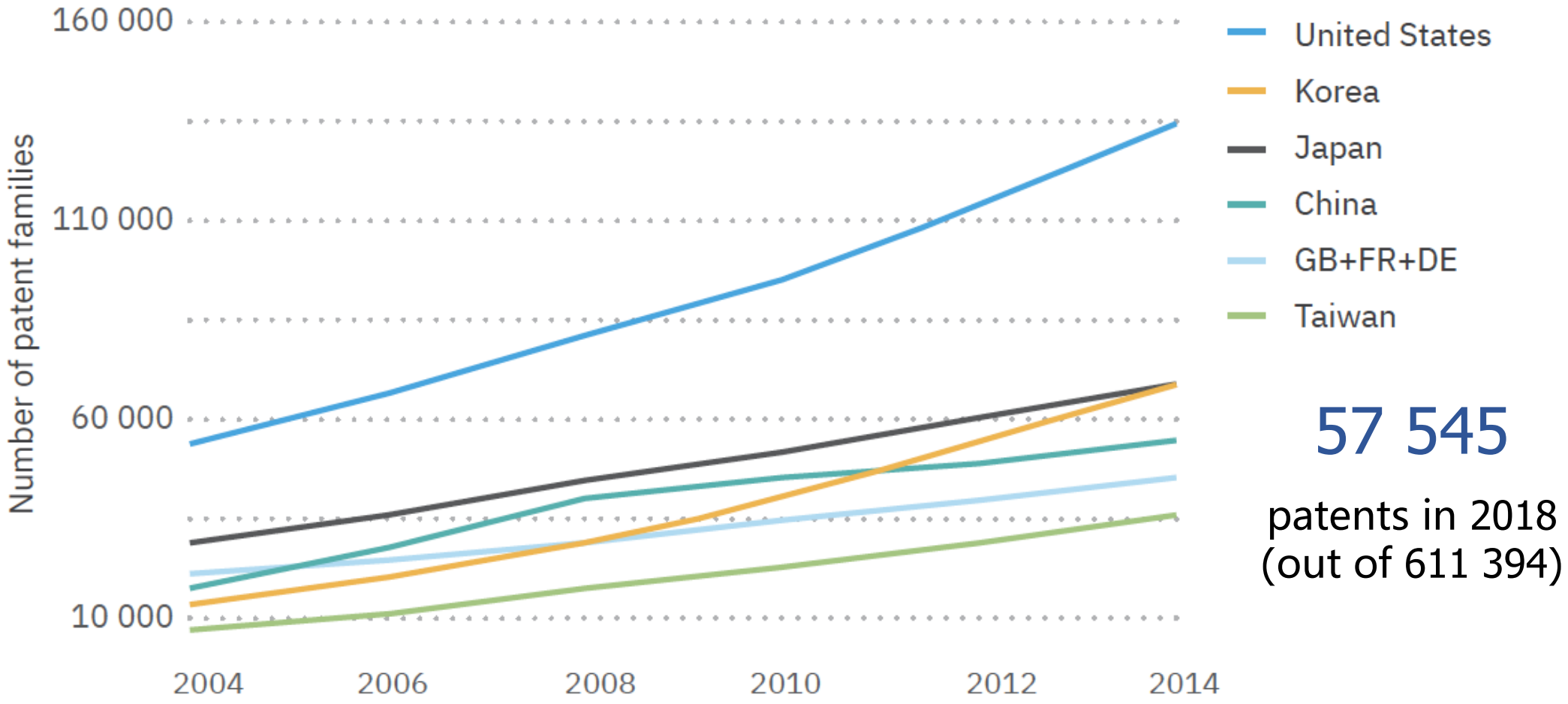


180 221

publications in 2018
(out of 4 866 023)

[Almanac AI #1’19]

AI Patents

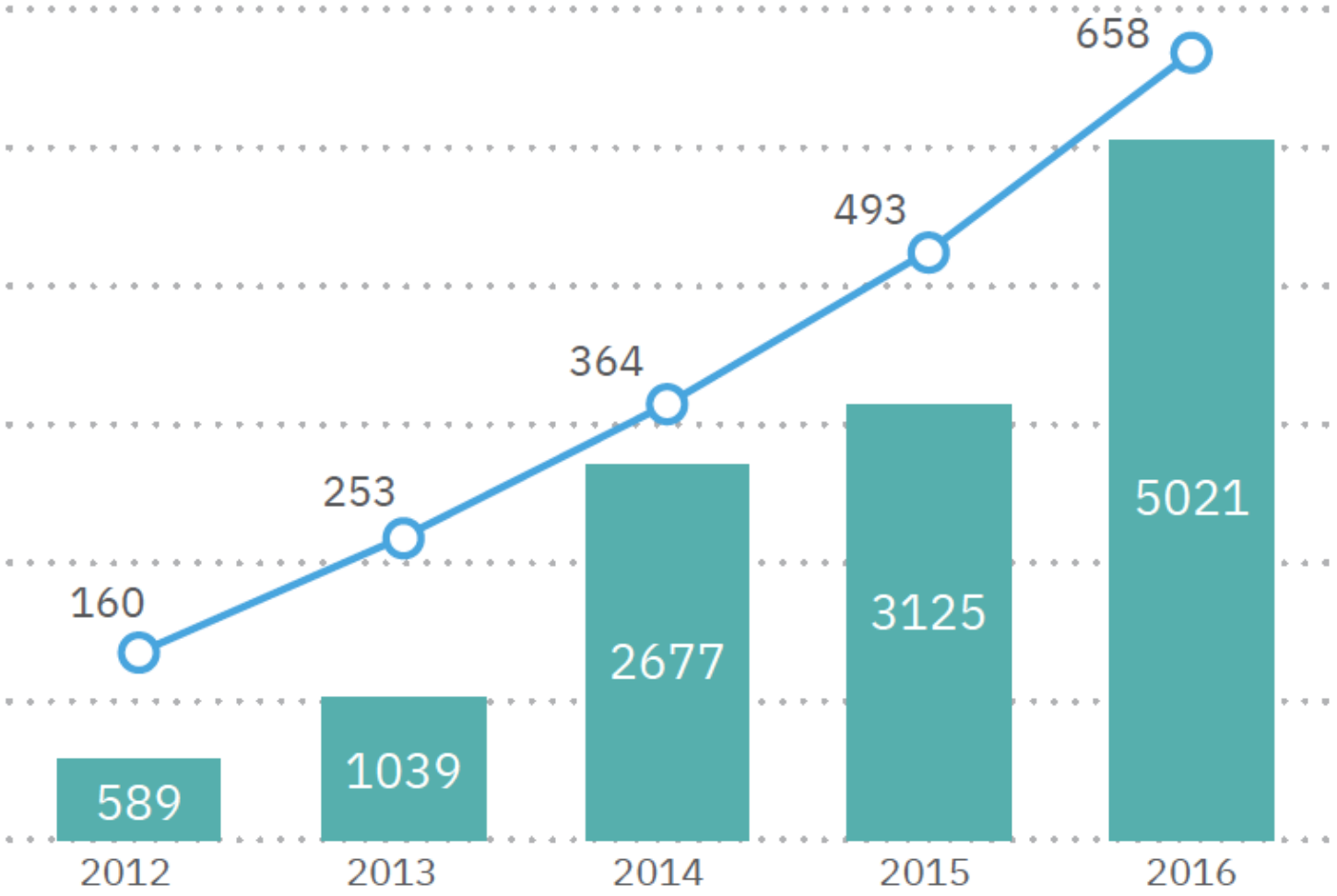


57 545

patents in 2018
(out of 611 394)

[Almanac AI #1'19]

Investments in AI projects



5 809

grants in 2018
(out of 126 122)

[Almanac AI #1’19]

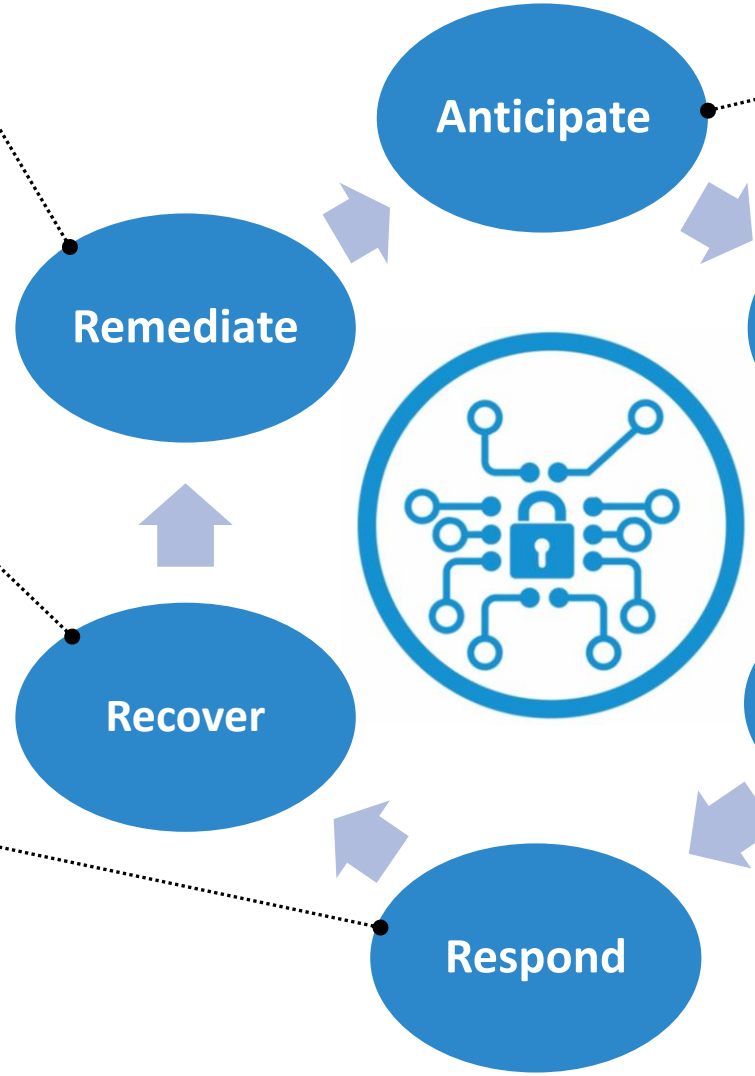
Cybersecurity

- Cybersecurity is the practice of protecting systems, networks, and programs from digital attacks. These cyberattacks are usually aimed at accessing, changing, or destroying sensitive information; extorting money from users; or interrupting normal business processes.
- Cybersecurity is the protection of computer systems from the theft of or damage to their hardware, software, or electronic data, as well as from the disruption or misdirection of the services they provide.
- Cybersecurity refers to the practice of ensuring the confidentiality, integrity, and availability (CIA) of information.

Some Types of Cybersecurity Threats

- Ransomware is a type of malicious software designed to extort money by blocking access to files or the computer system until the ransom is paid.
- Malware is a type of software designed to gain unauthorized access or to cause damage to a computer.
- Social engineering is a tactic that adversaries use to trick you into revealing sensitive information. Can be combined with any of the threats listed above to make you more likely to click on links, download malware, or trust a malicious source.
- Phishing is the practice of sending fraudulent emails that resemble emails from reputable sources. The aim is to steal sensitive data like credit card numbers and login information. One of the most common type of cyber attack.
- Denial of service attacks (DoS) are designed to make a machine or network resource unavailable to its intended users.

Full Cybersecurity Lifecycle



1

Security awareness
Security risk management
Vulnerability assessment
Penetration test
Governance, organization, process and procedure

2

Design and deploy of security architectures
Security system and device management

3

Managed security services
Monitor network data, digital identity
System/User behavioral analysis
Event correlation, log management
Malware analysis
Compliance control
Incident identification, classification and notification

4

Incident containment
Active defense, kick back
Support to crisis management process

5

Investigate and loss intrusion
Digital forensics
Incident eradication
Support to service recover

6

Design and deploy security posture improvement

Red and Blue AI

AI in cybersecurity is a set of capabilities that allows organizations to detect, predict, and respond to cyberthreats in real time using AI methods.

Red AI is AI based Cybercrime

Offensive AI, Adversarial AI

Cyberattacks which are using artificial intelligence, machine learning, and robotics in combination to make attacks even harder to detect and fight against.

Blue AI is Defensive AI

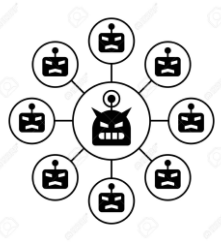
Blue AI seems to be the best defense against Red AI. Probably it is a new frontier of information warfare. The asymmetric, instantaneous nature of cyberattacks demand the adoption of autonomous defense systems that could act in response to an attack in an early stage.

This confrontation will probably generate an "AI arms race".

Red AI Applications

Malware creation

- Speed up creation
- Enhance evasive capabilities



Smart botnets

- Self-learning botnets

Phishing

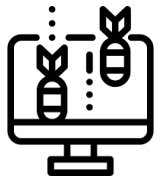
- Intelligent social engineering
- High quality auto-generated texts

Unauthorised access

- Breaking current CAPTCHA

Intelligent user modeling

- Classify victims
- Optimize algorithms



Feedback weaponization

- Poison ML to attack users with false alarms

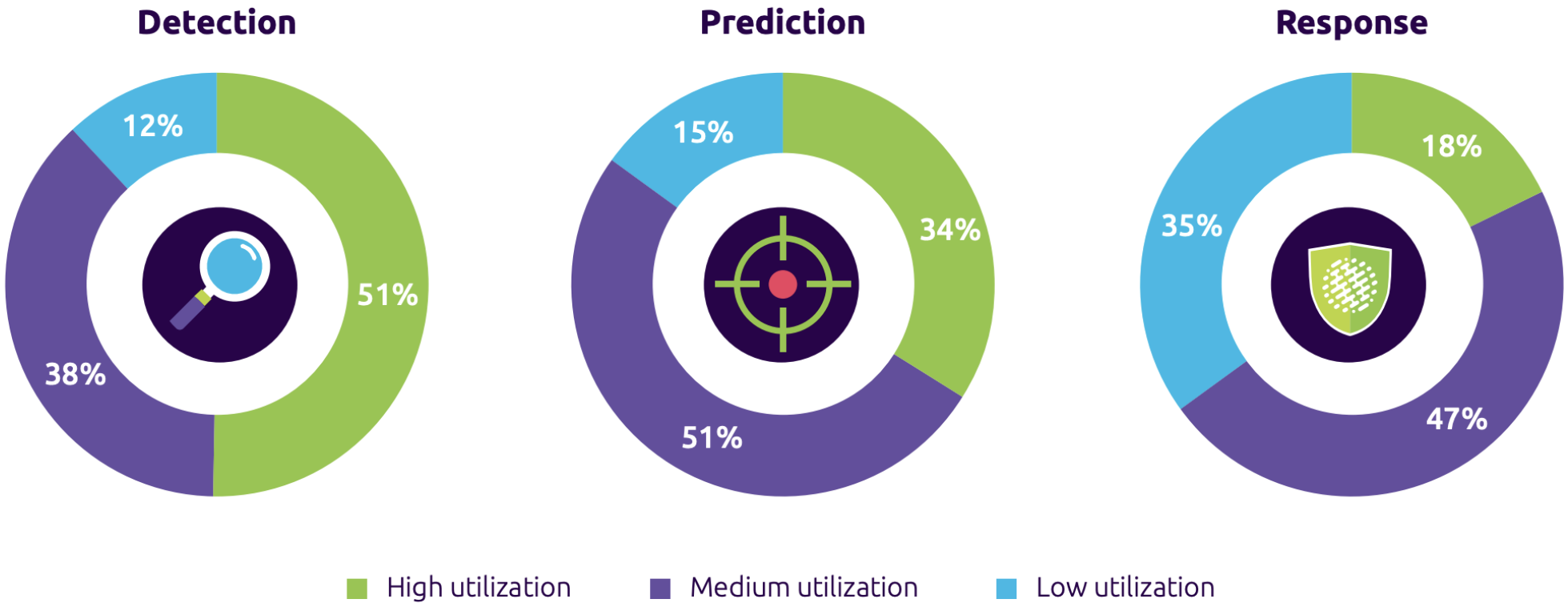
Data poisoning

- Feed poisoned training data to cybersecurity tools

Conditional attacks

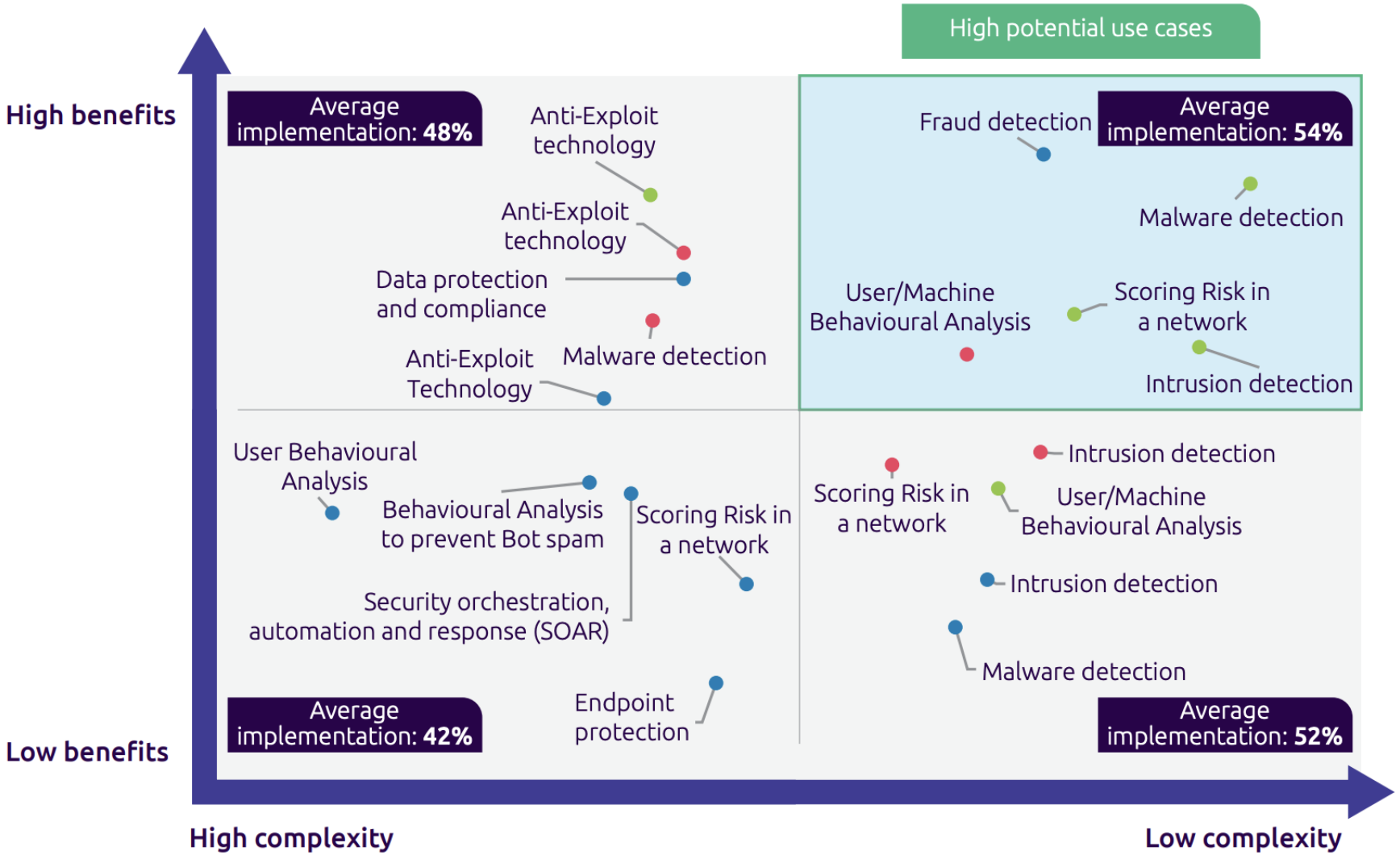
- Blockchain based smart contracts

Blue AI in Cybersecurity Capabilities for Business



[Capgemini Research Institute, AI in Cybersecurity executive survey]

Blue AI in Cybersecurity Use Cases



[Capgemini Research Institute, AI in Cybersecurity executive survey]

AI Methods for Cybersecurity Use Cases

High Potential Use Case	Description	AI Methods
Intrusion detection	Rapidly detect, analyze and defend against cyber attacks in real-time through automated, highly accurate insights into malicious activity.	Genetic algorithms to characterize the network traffic, and automatically generate signatures for blocking the attack. Deep neural networks to create models to detect attacks. Evolutionary computations to optimize detection strategies. Multi-agent technology to model intrusion process.
User/machine behavioral analysis	Identify behaviors that are unlikely to represent human actions.	Machine learning methods to detect new forms of cyberattacks in real time with high accuracy. Recommender system algorithms for detecting compromised accounts through suspicious user behavior. Genetic algorithms for user modeling.
Fraud detection	Detect possible fraud threats, reducing financial loss while also enhancing the user experience.	All kind of machine learning method to detect and predict fraud. Recommender system and social networks analysis methods to identify users in risk groups.
Malware detection	Use previously-identified characteristics of malware to predict potential future malware infections that signature based approaches may not be able to detect.	All kind of machine learning methods including deep learning to classify and/cluster malware software.
Scoring risk in the network	Compile risk ratings scores that are data-driven, quantitative, and not depend on cyber analysts insights.	Ontology to integrate security metrics. Graph analysis to optimize scoring process.

Research
of Laboratory of Computer Security Problems
SPIIRAS

Ontology of Metrics for Cybersecurity Assessment

Leader: Elena Doynikova.

Problem: number of different measures, tools, approaches, and huge stream of data to calculate security metrics for security assessment and countermeasure selection.

Goal: to process available data and knowledge to receive answers to the information security issues.

Approach: develop an ontology that will allow calculating security metrics that answer important security questions based on the relations between data sources, objects of security assessment, primary and integral security metrics.

Some security metrics

Asset Capacity	Network Resilience
Average Length of Attack Paths	Operational Capacity
Compromised Host Percentage	Resource Redundancy
Exploit Probability	Service Availability
Impact Factor	Shortest Attack Path
Number of Attack Paths	Severity Score
Network Preparedness	Vulnerable Host Percentage

[Cheng et al. Metrics of security]

Fragment of Security Metrics Ontology

Infrastructure objects:

- network,
- workstation,
- ...

Security information:

- product,
- configuration,
- weakness,
- attack,
- attacker,
- countermeasure,
- vulnerability,
- exploit,
-

Results

- Proposed ontology combines security data sources, security information, objects of security management subject area and security metrics.
- Metrics are represented as concepts. It allows using the inference mechanism for calculation of integral metrics on the basis of primary metrics.
- Advantages: extensibility in terms of metrics (i.e. we can easily add and link new metrics) and universality (i.e. the proposed ontology can be used for security assessment of systems of various types).
- The extensibility of the ontology allows us to create the complete system of interconnected known security metrics in the future.

Network Attack Detection System Based On Signature Analysis And Adaptive Classifiers

Leader: Alexander Branitskiy.

Problem: timely detection of attacks.

Goal: to detect the attacks using TCP/IP traffic analysis.

Approach involves:

- multilayer neural networks,
- networks with radial basis functions,
- Jordan's recurrent neural networks,
- fuzzy neural networks,
- support vector machines,
- combination of classifiers with a variety of low-level schemes and high-level classifiers;
- unique algorithm for hierarchical traversal of the classifiers tree, characterized by the presence of a “lazy” connection of individual classifiers and the ability to specify arbitrary nesting of classifiers.

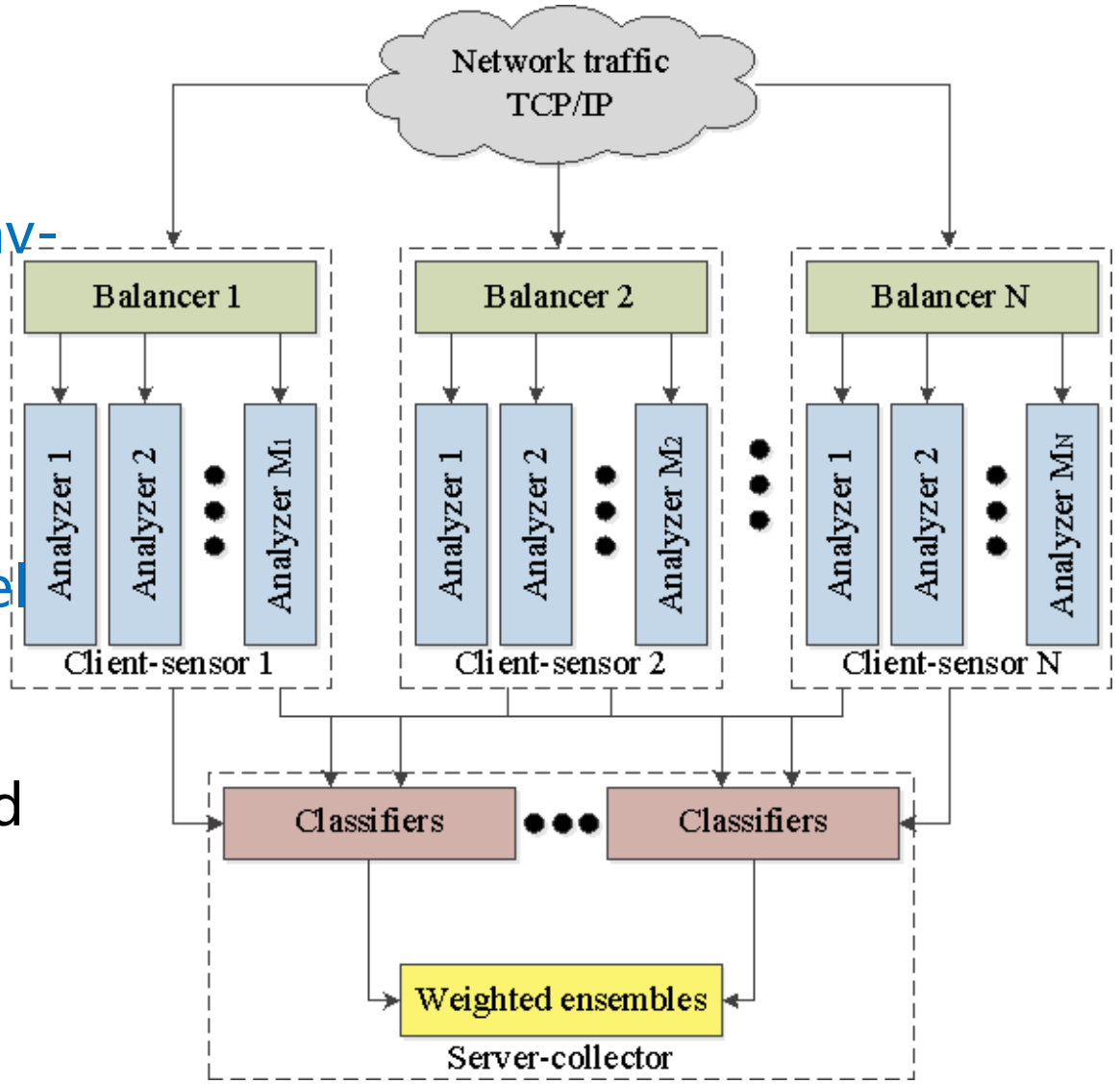
Network Attack Detection Prototype

Faster processing of network traffic due to the dynamic memory pre-allocation mechanism, fnv-hashing of network connections, built-in balancers, and an algorithm for bypassing multiple nested lists with network connection parameters.

A set of sensors with support for several parallel modifications of the substring pattern search algorithms.

Secure data channel between client sensors and server based on RPC protocol, encrypted using SSL.

Support for five AJAVUs for creating adaptive modules.



Advantages of Network Attack Detection System

The prototype allows to detect attacks with evading and attack with insertion (unlike analogues, e.g. Snort, Suricata).

It support runtime code editing.

It has lightweight interpreter for writing side scripts and defining the structure and functioning rules of adaptive classifiers and individual detectors.

20 plugins in the form of .so-libraries, which include various methods of computational intelligence and clustering, as well as schemes for their hybridization.

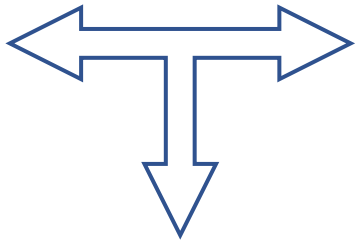
Other Applications:

- detection of statistically significant bursts in traffic parameters,
- malware detection,
- identification of attempts of unauthorized access to a computer network,
- detection of abnormal user activity and applications.

Intelligent system for detecting and countering inappropriate information on the Internet

Leader: Igor Saenko.
Problem:

The rapid spread of the Internet, including social networks, into the state, industrial, economic and socio-cultural spheres



Internet is becoming one of the most important threats to personal, public and state information security

Protection of individuals, society and the state from information that spreads through Internet and other networks and is capable of harming the health of citizens or motivating them to unlawful behavior is a problem of national importance affecting the provision of information security of the country

Why the problem is not solved good enough?

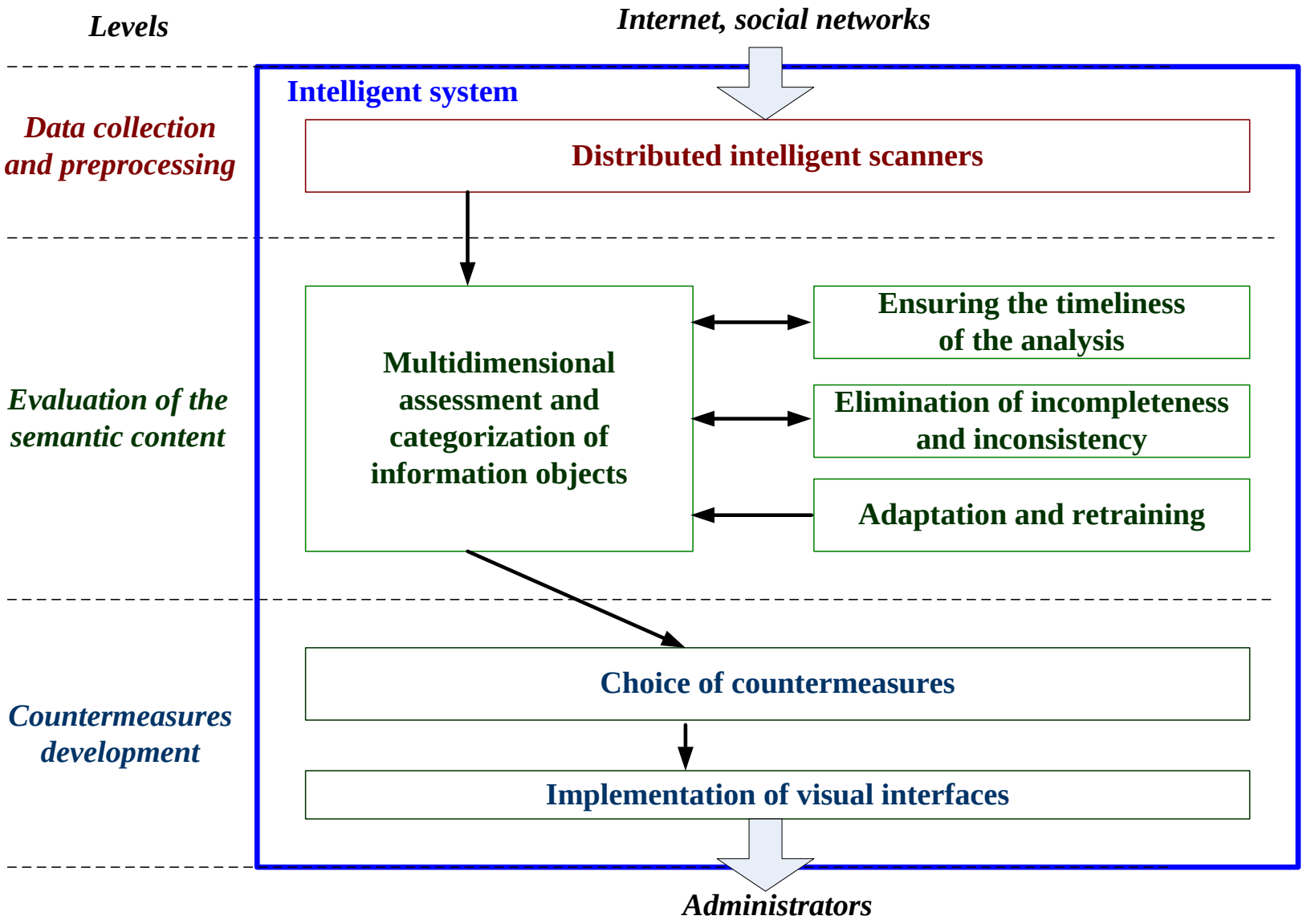
The known automated means of identifying and counteracting inappropriate information on the Internet do not meet the requirements for:

- speed,
- completeness,
- accuracy,
- adequacy of the decisions made.

Reasons:

1. Web resources have a complex hierarchical structure and consist of many disparate elements.
2. It is necessary to process big data online.
3. The analysis of information on the Internet is associated with the processing of conflicting and changeable data.

The architecture of the system



Implementation of the Classification Component

Classification component solves the problem of automatic classification of network information objects using data collected in a distributed storage using distributed intelligent scanners.

Any combination of the available characteristics of information objects can be considered.

Natural language text is the most important.

Training sample is a set of pre-categorized information objects, for example, webpages.

Research results

- The work offers a new approach to creating an intelligent system for detecting and countering inappropriate information on the Internet based on the use of machine learning methods and big data processing.
- The proposed system architecture contains 7 components .
- Experimental evaluation of the developed component of multidimensional evaluation and categorization of information objects in single-threaded and multi-threaded modes has shown the high efficiency of application of various classifiers.

Future research is associated with the development, integration and evaluation of all system components on a much larger dataset and in a more productive computing environment.

Conclusion

A constantly evolving IT landscape and ever-expanding attack surfaces lead to increasingly complex security challenges.

AI and machine learning are redefining every aspect of cybersecurity today.

Cybersecurity applications are among the most popular AI applications today.

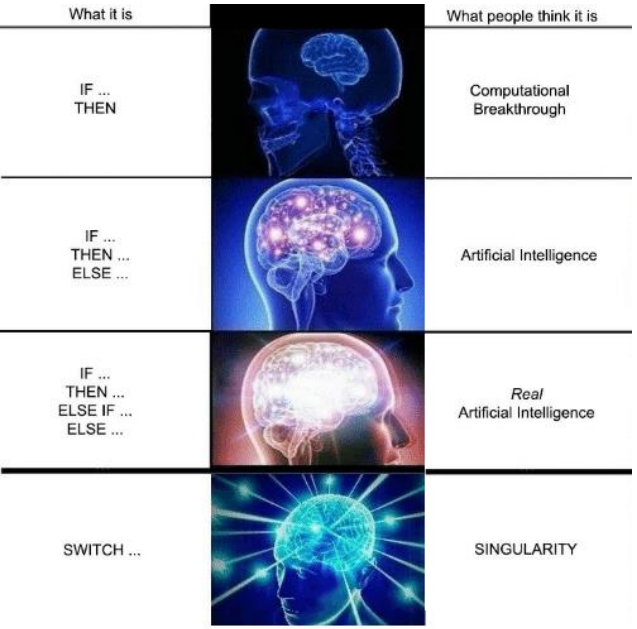
As AI improve, it will be leveraged by malware authors and cyberattackers for a variety of different applications including network scanning, automated phishing attacks and AI-enabled botnets.

AI-based defenses will have to learn to adapt and grow to meet the threat of these rapidly-evolving systems as well.

AI is the future of cybersecurity.

Thank you for attention!

Questions?



Dont even get me started on Deep learning

Contacts:

Olga Tushkanova
tushkanova.on@gmail.com