Cyber Week

June 27th-30th, 2022

Tel Aviv University, Israel

Yuval Ne'eman Workshop for Science, Technology and Security Tel Aviv University

ICRC Blavatnik Interdisciplinary Cyber Research Center

TEL AVIV UNIVERSITY אוניברסיטת תל-אביב

Cyber Israel National Cyber Directorate

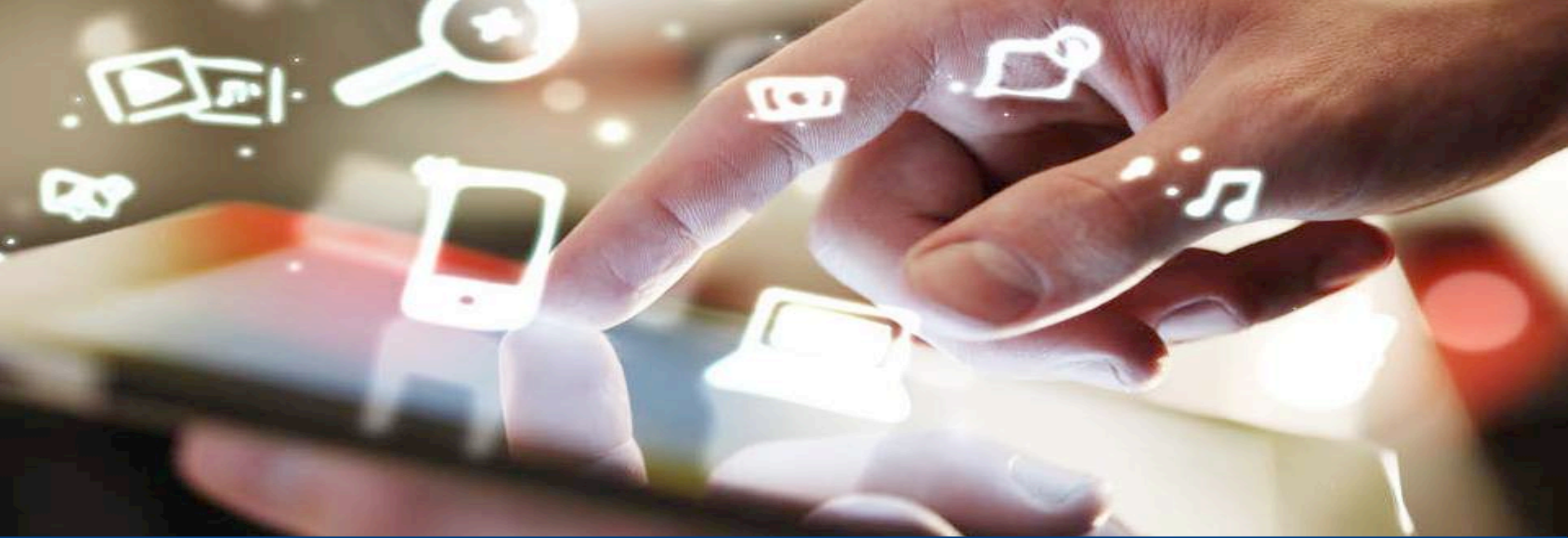
In cooperation with:

Ministry of Foreign Affairs Israel

ISRAEL CYBER ALLIANCE

State of Israel Ministry of Economy and Industry Foreign Trade Administration

ISRAEL EXPORT INSTITUTE



A Business Transformation Journey via Disruptive Cyber Protection Technologies

by Jean Lehmann, CEO - Cyber Capital HQ   www.cybercapitalhq.com (ISO27001 Certified)

CCHQ
Creative Cyber & Data Solutions

Cyber Week
June 27th-30th, 2022
Tel Aviv University, Israel

ICRC Blavatnik Interdisciplinary Cyber Research Center

TEL AVIV UNIVERSITY

Cyber Israel
National Cyber Directorate

Artificial Intelligence and Cybersecurity

AGENDA

I – Cybersecurity challenges and how AI can help

II – Use cases of AI applied to Cybersecurity

III – Future trends and developments in AI and Cybersecurity

Introduction and Purpose

Artificial Intelligence and Cyber Security

This presentation will explore the **next frontier of AI applied to Cybersecurity**.

In the first part, we will present various AI use cases of **the intersection between Cybersecurity and Artificial Intelligence**, such as preventing zero-day threats. We will also introduce other techniques and approaches, such as behavioral techniques based on machine learning to identify deviation from a baseline norm to recognize and contain cyberattacks.

In the second part, we will look at possible future trends of how we can apply AI to Cybersecurity by optimizing capacity to resources ratios and making the relationship between people, processes, and technology more efficient.

We will also touch upon **the convergence of Cybersecurity with AI Ops and Business Intelligence** and look at how we can apply those techniques for **Cybersecurity predictive analytics** in the context of Secure Operations Centers in identifying where potential threats are likely to emerge from within or outside an environment.

Three waves of AI

CCHQ
Creative Cyber & Data Solutions

DARPA (defense Advanced Research Projects Agency) defines three waves of AI

Wave	Description	Characteristics	Examples
Handcrafted Knowledge	Rule-based automatic application of rules for logical reasoning – defined patterns	• No learning applied to higher levels • Poor handling of uncertainty • High false positive rates	**DARPA** Cyber Grand Challenge (automated software vulnerability analysis and repair)
Statistical Learning	Machine learning to perform probabilistic decision making.	Classification and prediction but lack of context understanding – minimal reasoning ability	ML models (supervised/unsupervised/semi-supervised) have fundamental advantages compared to signature-based AV, such as, identifying zero-day attacks and targeted malware
Contextual Adaptation	• Models to explain decisions • Models to drive decisions	Perceiving Learning Abstracting Reasoning	

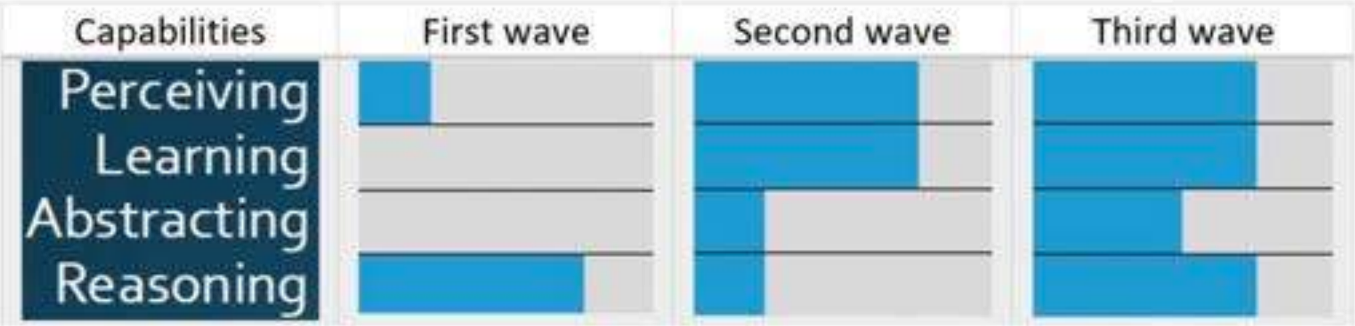
Three waves of AI

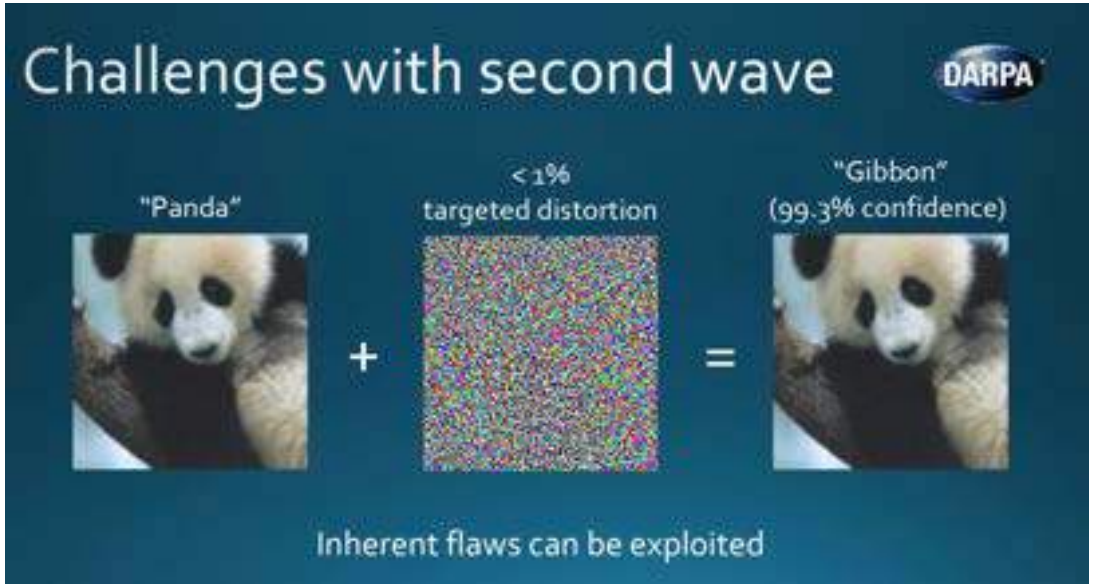
The **first wave of AI** is "handcrafted knowledge". Logical reasoning can be implemented by taking the knowledge about a particular domain and characterizing it in set of the rules that could be fit in the computer. The computer then studies the implication of those rules.

The **second wave of AI** is "statistical learning" which is good at perceiving the surrounding world, for example distinguishing between voices and faces. Powerful in classification and prediction when provided with the context however limited capability in understanding the context and minimal reasoning abilities too.

The **third wave of AI** relates to "contextual adaptation". In this wave the system itself will build over time the underlying explanatory models for classes and real-world phenomena. The third wave provides further potential for explainability, visibility, and transparency into how AI algorithms make decisions, predictions, recommendations, and classifications.

AI is a technology, regulatory, and policy issue with many ramifications on ethical and legal challenges (issues of transparency, trust, interpretability, explainability, visibility, actionability).





Concepts

CCHQ
Creative Cyber & Data Solutions

Machine Learning Algorithms

- **Training:** model learns from the data set of known samples
  - More intense computational operation (deep neural network: months to train large clusters
  - Labels on data sets (malicious/nonmalicious)
- **Prediction:** trained model makes an educated guess about a new, unknown sample
  - Prediction follows training, still requires significant memory and CPU usage.
  - Distributed training locally: benefits and challenges-performance

Runtime: environment where training/production could occur (local/endpoint or remote/cloud)

Features

- File size
- Entropy
- Derived features based on other features
- Information exposure and encoding

Feature Engineering - What and how information is exposed - How to encode information

Disadvantages in using too many features: risk of overfitting, higher resource consumption, vulnerabilities. Interpretability, efficacy and robustness of a model depends on features.

Data sets

Data set to train the model needs to be representative of the real world

Evaluating the data source for the degree of trust and reliability

Label noise (mislabeling) can bias the model (bias may be coming from data or algorithm)

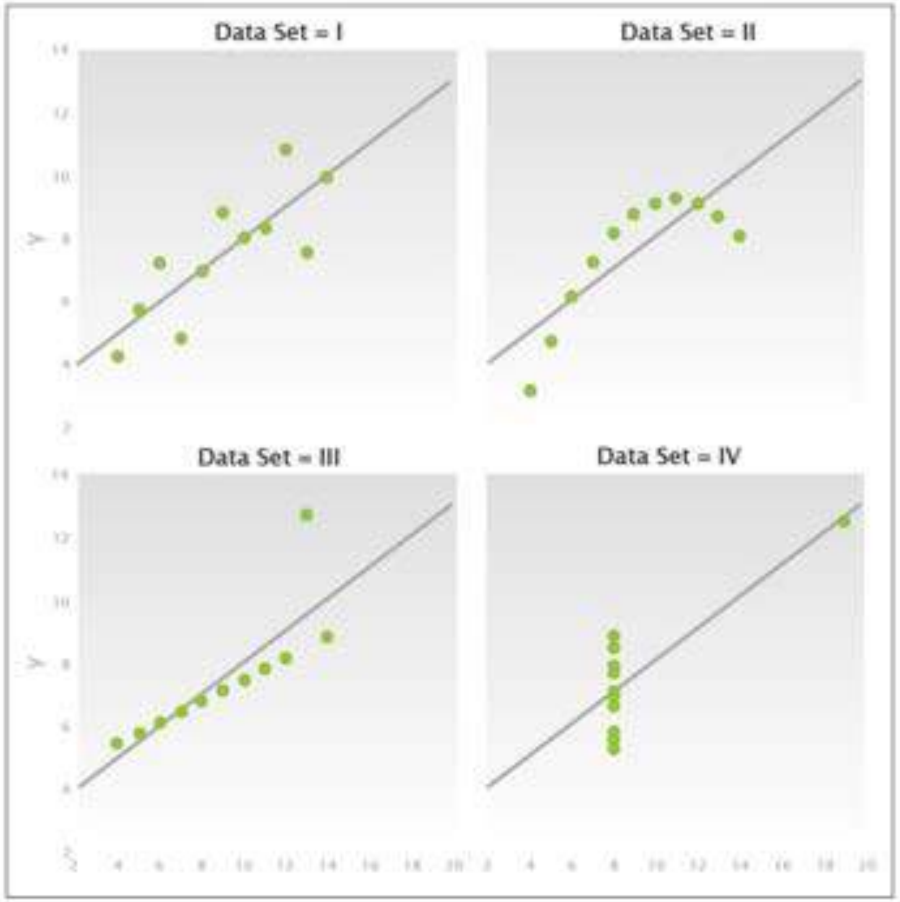
Risk of unbalanced data when one label occurs more frequently than others. Data sets and training sets should be representative of the diversity to mitigate the risk of bias.

Continuous monitoring of the data for consistency

Crucial feature pre-processing (normalization, weighting schemes)

Evaluating the data source for the degree of trust and reliability

Huge data set does not guarantee performance. A good data set shows variety and fair representation



Anscombe's quartet

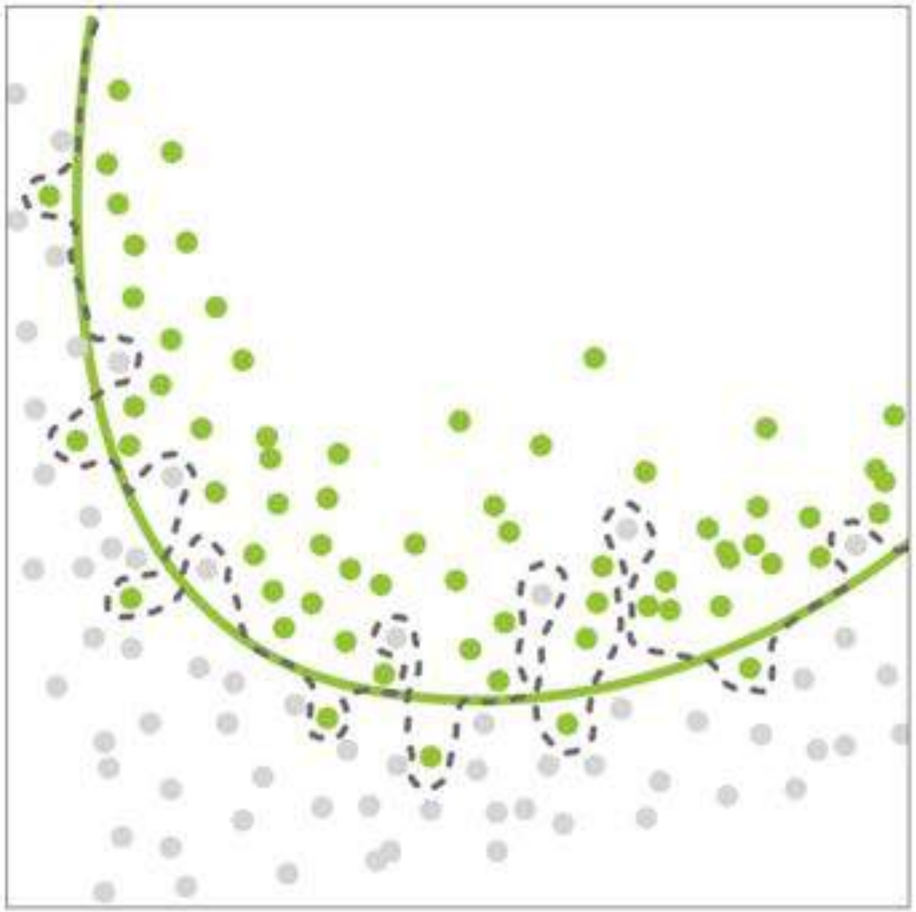
Different data sets and same predictive model

Only upper-left model better fits its data set

Integrating human validation, tools for understanding the model – Explainability to validate the data

Tools for explainability may expose vulnerabilities and cause IP Leakage

Goodness of fit



Underfit ⟺ oversimplified – generalization – poor efficacy

Overfit ⟺ Too specific – does not transfer well to new samples in the real world

Dotted line is an overfitted model

Green line: appropriate decision boundary – Performance will be better for new points

Definition of Generations

- Cybersecurity machine learning generations depend on five main factors that are reflective of the intersection of AI/Data Science and Cybersecurity. The majority of technologies that integrate machine learning will be in the first or second generation category.

Runtime	The location where the ML training and prediction occur (in the cloud or locally in the endpoint)
Features	Feature engineering – how many features are generated? How are they pre-processed and evaluated?
Datasets	How is trust handled in the process of data curation? How are labels generated, sourced, and validated?
Human Interaction	How do people provide feedback and understand the model decisions? How are models monitored and overseen?
Goodness of Fit	How well does the model reflect the datasets? How often does it need to be updated?

Generations of Machine Learning

Generation	Runtime	Features	Data Sets	Interactivity	Goodness of Fit
First	• Cloud training • Cloud prediction	• Over 1000 features	• Over 1M data examples • Human labeled	• Human understands decision	• Underfit • High false positive rates
Second	• First generation • Local prediction	• Over 100,000 features	• Over 100M data examples • Human/heuristic labels	• Model struggles to explain decisions	• Overfit, misleading false positive rates
Third	• Second generation • Cloud enhanced models	• 1 to 3M features	• Over 1B data examples • Largely heuristic labeled	• Model provides explanations understandable	• Fits appropriately, accuracy metrics generalizes
Fourth	• Third generation • Local training	• Over 3M features	• Online learning	• Model explains strategy, high-level feedback	• Model fits current as well as future inputs
Fifth	• Fourth generation • Unsupervised local training	• Unlimited with semi-supervised discovery	• Active learning	• Human input optional, interpretable insights	• Model identifies and adapts to concept drift

First Generation: the feature set is small. High rate of false positives. Limited efficacy. Easy to bypass. Cannot be deployed on endpoints

Second Generation: heuristics are used to supplement human labels. Allows for local model predictions, still requires cloud based training. Models are overfit to training data. Some predictive power, periodic updating to avoid concept drift.

Third generation: cloud model complements the local model. Models designed to be hardened against attacks. Concept drift mitigated by better fitting and generalizability.

Fourth generation: features designed by strategic interactions human-models. Human feedback to correct the model.

Distributed, semi-supervised learning. Human analysis guided by model-provided insights.

Real-Time business Transparency Applications to Cyber and SOCs

Security remediation roadmaps and RSI as potential outputs

Threat Intelligence

Asset inventory and visibility

Security posture monitoring

Identity and Access Management AD/Auth building access

HR ERP

XDR Endpoint Monitoring

Agent

Attack Surface Monitoring

→

Risk CIA: Confidentiality, Integrity, Availability

Impact assessment: market, financial, reputational, operational, strategic, customer

Probability of "it" happening

Definition of "it"

Measurement of network strength in $, time, competencies and associated investment

Create capital or knowledge intensity to limit risk

→

Process mining

Recommendations

Automation

Alerts

Signals

Predictions

Orchestration

Knowledge Graph

Effort model (people, category, metrics)

→

Secure Operations Centre

Board – defined reporting

Compliance/Regulator - "legal" documents

Third-parties as required

Repetitive work:

- Eliminate unnecessary work for increased efficiency and lower operational risk:
  - Reduced ticket volume
  - Automate the right process
  - Uncover patterns of incoming work

Service desk capacity:

- Effectively match capacity with demand to improve service levels at the right cost:
  - Reduce staffing costs
  - Increase service levels
  - Accelerate the scheduling process

Ticket Category Metrics:

- Measure efficiency and service equality by agent:
  - Reduce mean time to resolve MTTR
  - Reduce escalation rates
  - Increase employee performance

Process Mining:

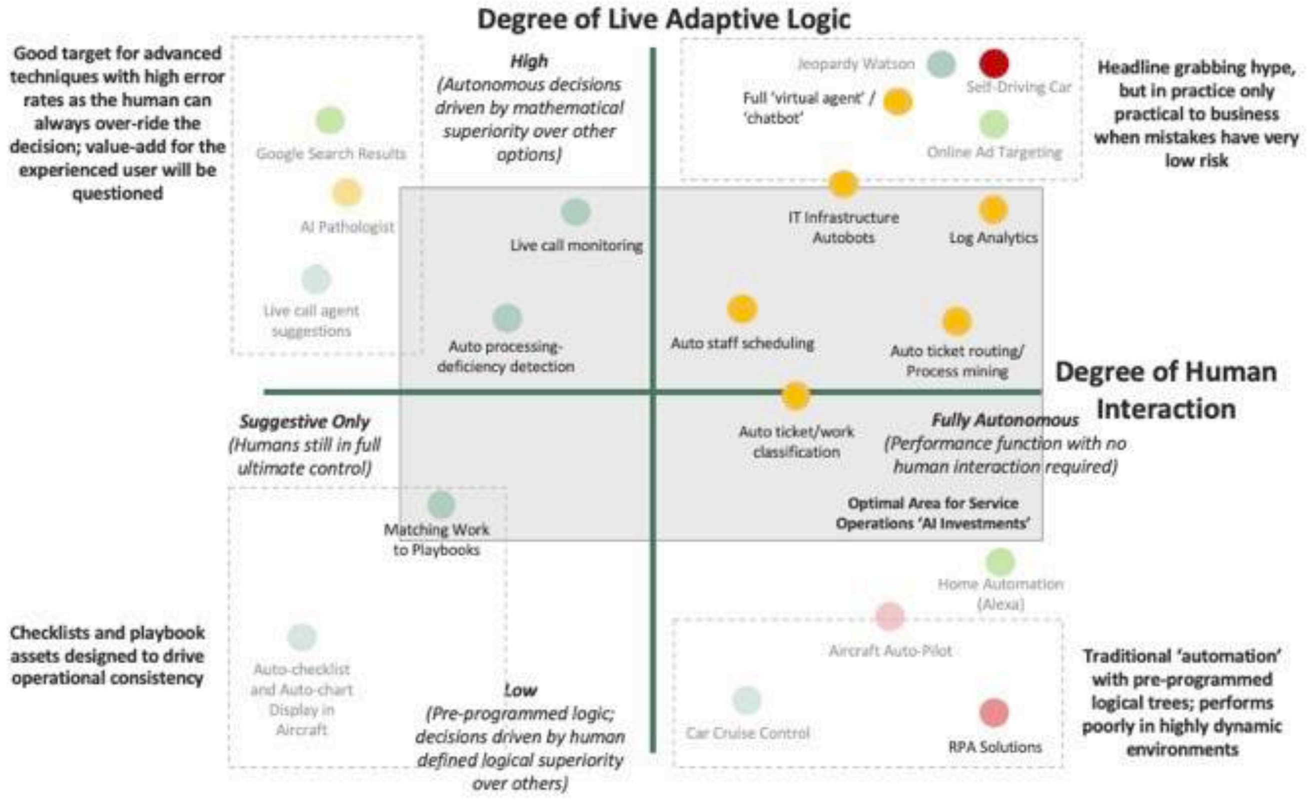
- Streamline processes by uncovering patterns of wasted time and effort
  - Eliminate bottlenecks
  - Eliminate ping-pong
  - Reduce number of agents touching tickets

→ **Automation** → **Eliminate** / **Optimize** / **Automate**

Automating often fails to return its expected value. Automating a process that could be removed doesn't improve performance and efficiency.

Automation works best on optimized processes that are consistent and repeatable.

Human interaction relative to Live adaptive logic



Impact of a Bad AI Decision

Low (Minor annoyance, easy to recover, etc.) — *High (Death, major business impact, etc.)*

Where to target AI investments:

- Fully autonomous AI technologies grab headlines, but are practical only where the risk of errors (which do happen) is low
- While less heading grabbing, significant value can be gained by operations where AI investments seek to not fully replace humans but rather empower them to make better decisions based on a large body of historical knowledge
- Ultimately, starting with a defined business problem, breaking that down into specific constituent AI/ML components and integrating that into operations with risk in mind will yield the best business value
- Many investments can and should be complementary. For example, while many wouldn't consider most automation or RPA to be 'AI', auto process defect assets are key tools to correctly targeting automation investments

Three core forces of service operations

Service: the quality of the operations output, often indirectly quantified by service level agreements, customer satisfaction or other requirement.

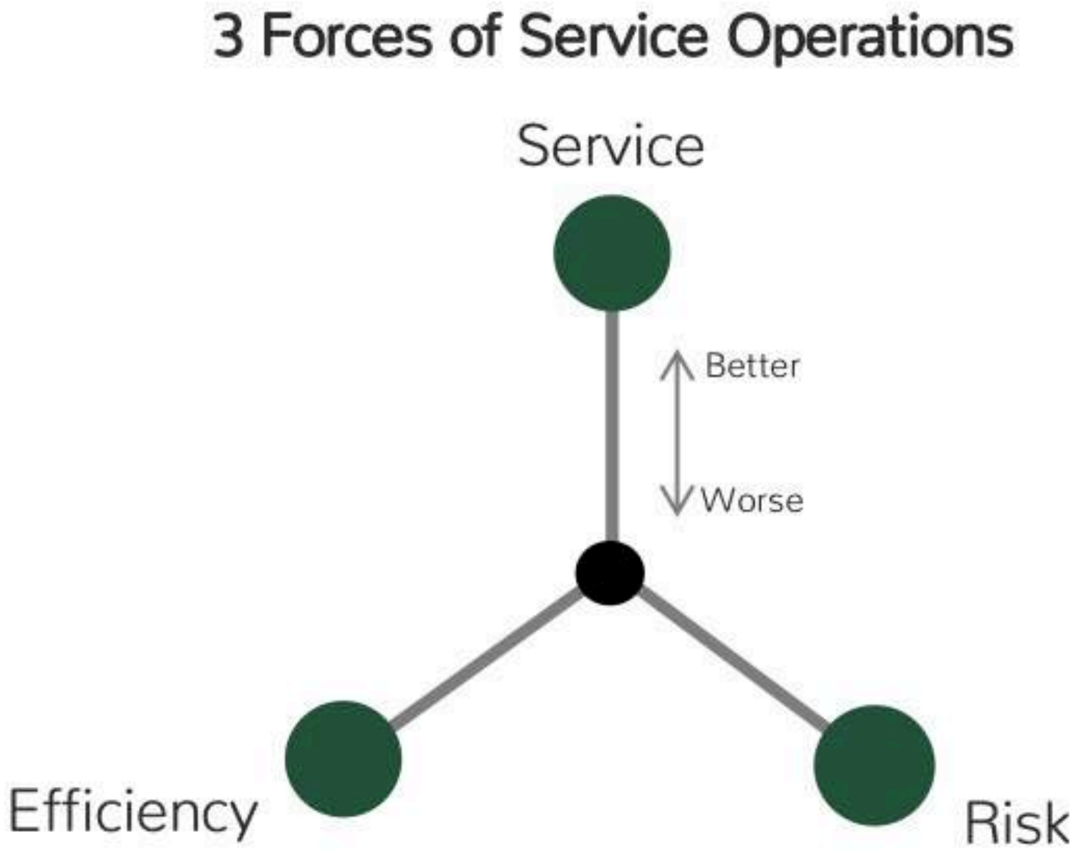
Efficiency: relative amount of resources required to complete an operation, typically correlated with cost

Risk: the potential of uncontrolled negative impacts on the operation.

The 3 forces of service operations tend to have a reflexive negative impact on each other.

Modern data-driven business leaders will increasingly demand their data to provide continuous real-time insight into these three forces, with those insights presented in an actionable context that allows to act upon a situation wondering off course. AI + Human beats AI only at chess and the same applies in operations centres.

Data should be used not only to monitor the impact of previous actions but also to plan and project the impact of future actions. Real-time analytics data can help optimize capacity to resources ratios. Respecting the three forces when managing, analysing and reporting in complex service operations provides a strong foundation for balancing the needs of diverse stakeholders while promoting long-term sustaining business growth.



Cyber Capital HQ

CCHQ
Creative Cyber & Data Solutions

Jean Lehmann

CEO & Founder

jean.lehmann@cybercapitalhq.com

Jean Lehmann is a thought leader in innovative and disruptive technologies. Jean has particular expertise in Digital Transformation, Artificial Intelligence, Cyber Security, Blockchain, the Hedge Fund Industry. He has 20 years of experience in leading complex and multi-disciplinary engagements for leading companies. He is a regular contributor, panelist, and moderator at renown international Cyber Security conferences. He is an advisory board member in the Blockchain and Venture Capital sectors, a guest lecturer at INSEEC Business School on Hedge Funds and leads a Cybersecurity course for Executives at Paris Dauphine University. He is a trusted advisor to Arie Capital Banking Group, the first specialized online digital corporate banking group. Jean holds an INSEAD/Wharton MBA, a DEA from HEC School of Management, an MSc from Eurecom/EPFL, and an Executive Certificate in Public Policy from Harvard Kennedy School covering Cybersecurity: The Intersection of Policy and Technology, Artificial Intelligence, and Negotiation Strategies. He is fluent in English, French, German and Portuguese.