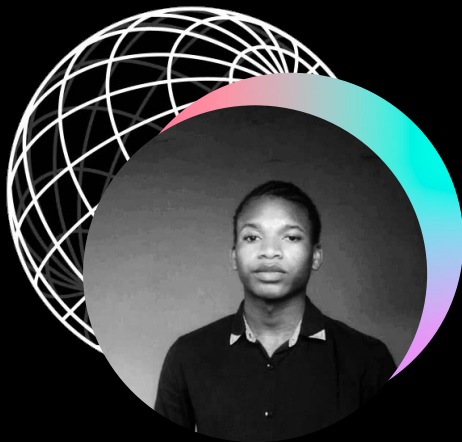


MOZDEVZ 10

# Machine Learning na Segurança Cibernética

Cyber  
Talk

# About 0x73



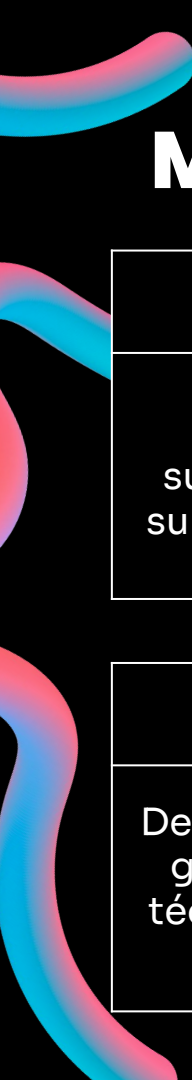
**Arlindo Cossa  
Júnior**

- Cyber Security Leader with 8+ years of experience and with clients across multiple industries
- Specialized in cyber security strategy, business resilience, cyber defense, cloud protection, incident response and managed security services.
- Ex Head of Security @ ATHSec
- Security Technology at Mining Industry
- Microsoft Security Architect / Researcher



<https://www.linkedin.com/in/arlindo0x73/>

**Cyber  
Talk**



# ML e AI – O que é?

## **Machine Learning**

Extraí padrões dos dados, com supervisão em dados rotulados e sem supervisão em padrões independentes.

## **Deep Learning**

Redes neurais complexas extraem características de dados brutos

## **Data Mining:**

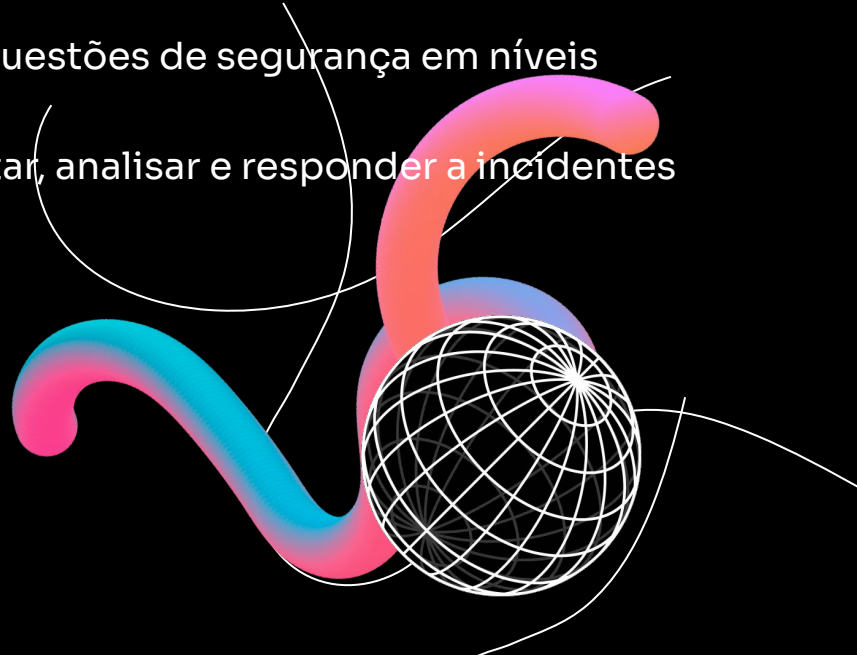
Descobre insights automaticamente em grandes conjuntos de dados, usando técnicas como agrupamento e redução de dimensionalidade.

## **Artificial Intelligence:**

Engloba a inteligência das máquinas, incluindo robótica e processamento de linguagem natural, buscando imitar o raciocínio humano.

# SOC (Centro de Operações de Segurança)

- O SOC é responsável por monitorar e gerenciar a postura de segurança de uma organização.
- É uma unidade centralizada que lida com questões de segurança em níveis organizacionais e técnicos.
- A equipe do SOC é responsável por detectar, analisar e responder a incidentes de cibersegurança.



# Questões Comuns

1. Muita informação para poucos operadores de sala de controle

2. Operadores sobrecarregados por sinais: eventos, estados, diagnósticos, alarmes, avisos, etc.

2.1 Software utilizado: SIEM (Gerenciamento de Eventos e Informações de Segurança)

2.2 Muitos alarmes falsos e alarmes de incômodo

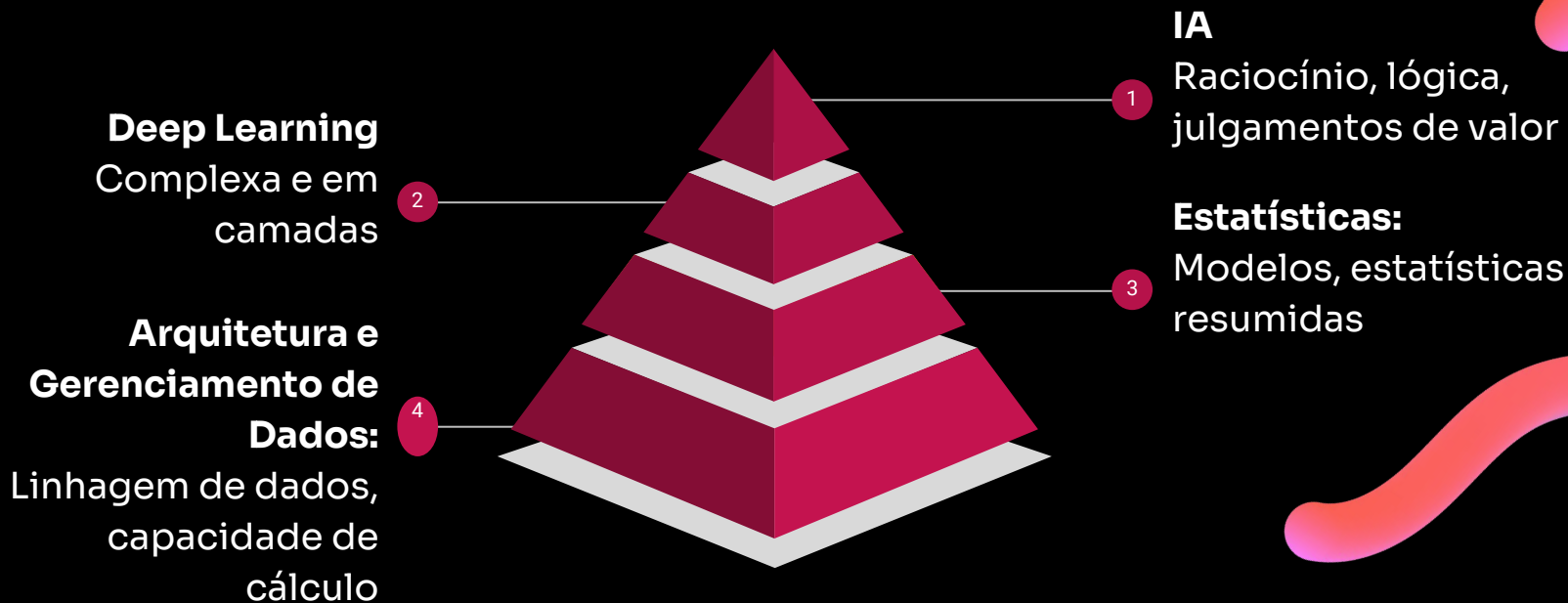
# Fusão de Informação – Coleta e Processamento de Dados.

## Como Funciona?

Os dados são coletados de várias fontes, incluindo dispositivos de rede, servidores e aplicativos.

Os dados são então processados e analisados para identificar ameaças de segurança potenciais.

# Pirâmide de Complexidade e Inteligência em Análise



# Riscos no Desenvolvimento Analítico

Inteligência deficiente leva a decisões comerciais erradas

Clientes insatisfeitos, ROI e ROA reduzidos

Falta de crescimento e geração de caixa

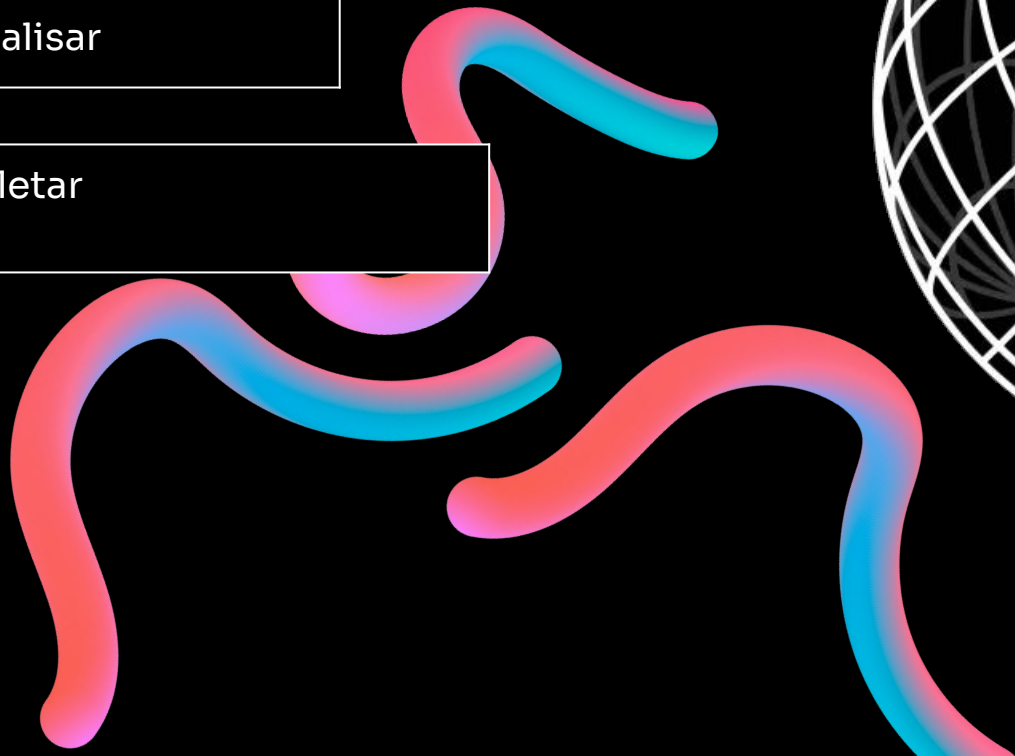
Aumento de falsos positivos e falsos negativos



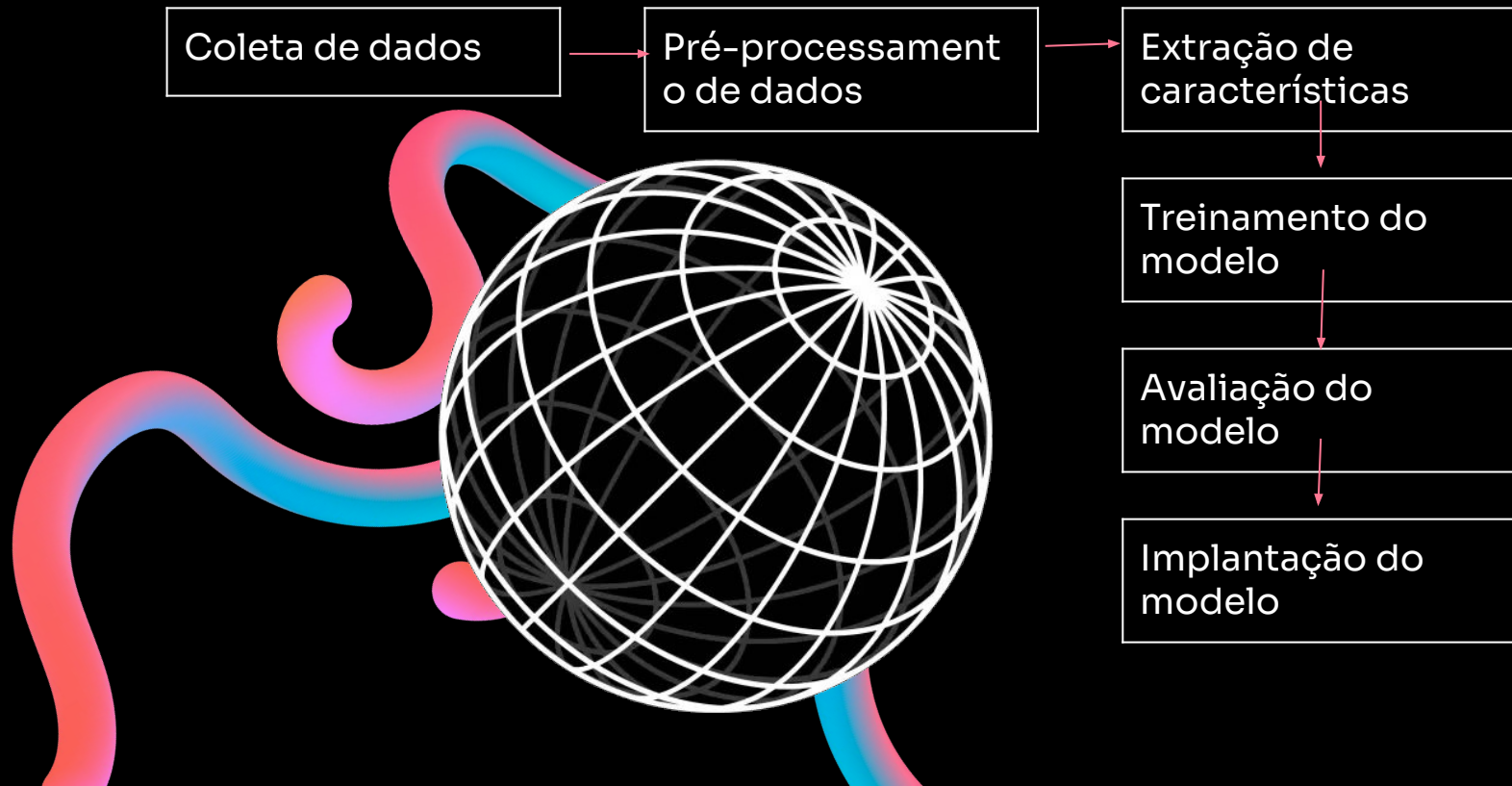
# Análise de Ameaças

Analisar

coletar



# O Processo de Aprendizagem



# Usos de Aprendizagem Automática em Segurança

## **Aprendizagem Supervisionada:**

- Classificação de Malware
- Identificação de Spam
- Análise de Dados de Firewall

## **Aprendizagem Não Supervisionada:**

- Análise de DNS
- Criação de Feed de Inteligência de Ameaças
- Automação de Analistas de Nível 1
- Análise de Comportamento de Usuários e Entidades



# Encontrando anomalias



## Dados fornecidos:

Comunicações de redes (ou seja, redes, sistemas, app's, etc..)

## Tarefa:

encontrar anomalias/ataques

# Deep Learning – as soluções para tudo

## promete algumas coisas

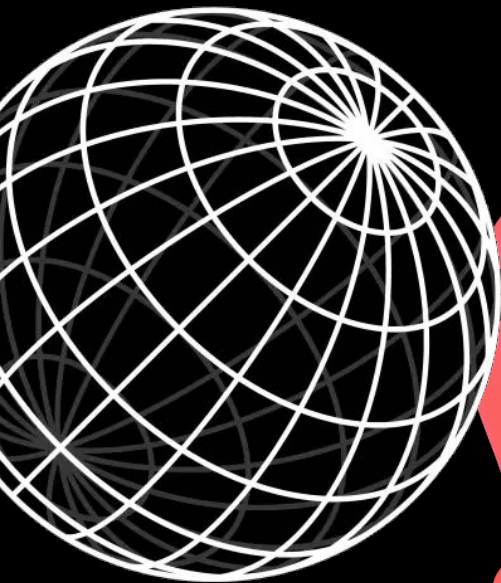
- Extração automática de características
- ALTA PRECISÃO DE DETECÇÕES

## requisitos:

Muitos dados disponíveis Mas:

- um único registro não indica o estado: bom/ruim
- não há informações suficientes dentro dos fluxos - precisa de contexto
- não há rótulos

# Agrupamento de tráfego para encontrar anomalias



- Limpar os dados
- Criar funções de distância
- Descobrir o algoritmo certo
- Aplicar os parâmetros algorítmicos corretos
- Interpretação dos dados

# Construindo Modelos de Aprendizagem Automática com foco

- Qual é o nosso objetivo?
- Que comportamentos podemos observar?
- Quais são os factores observáveis que reduzem a incerteza da interferência central do dispositivo comprometido?
- **As observações devem ser consistentes em todos os clientes e ambientes, evitando dependências locais.**
- Temos esses dados?
- Precisamos de contexto para isso?

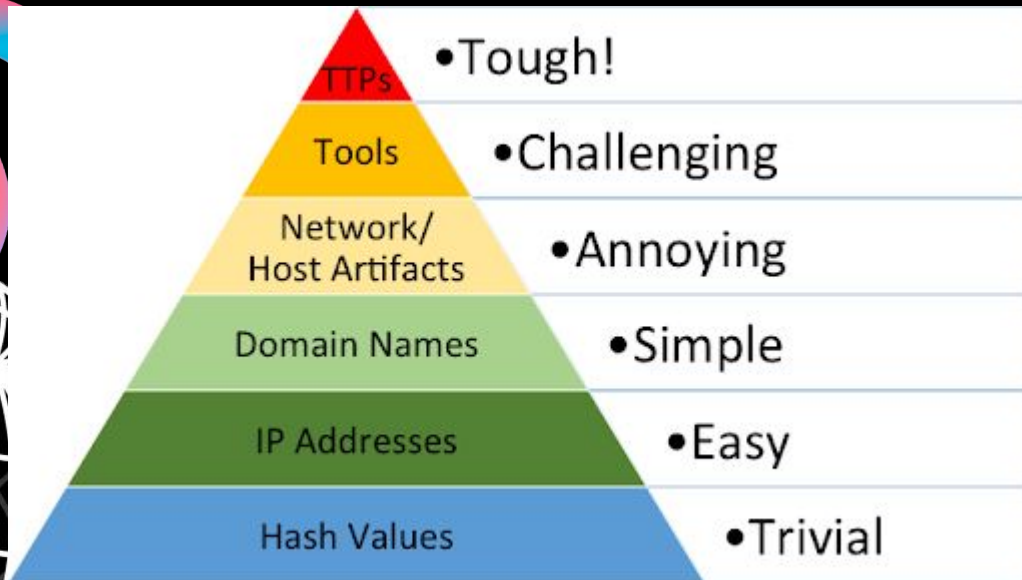




# Precisamos observar atividades nas seguintes áreas:

- Tráfego de rede
- Logs do sistema

- Comportamento do usuário
- Comportamento do aplicativo
- Infraestrutura em nuvem





# Quais são os factores observáveis que reduzem a incerteza da interferência central do dispositivo comprometido?

## **Estado de Hosts Suspeitos:**

- execução de processos incomuns ou a presença de portas abertas inesperadas.

## **Indícios de Inteligência de Ameaças sobre Comprometimento:**

- Utilizar feeds de inteligência de ameaças para identificar hosts que foram comprometidos.

# Quais são os factores observáveis que reduzem a incerteza da interferência central do dispositivo comprometido?

## Utilização Suspeita de Protocolos:

- Identificar hosts dentro de uma rede que estão a utilizar protocolos suspeitos, como aqueles utilizados para acesso remoto ou transferência de ficheiros.

## Túneis de Dados de Devices:

- Identificar hosts dentro de uma rede que estão a encaminhar dados através de outros protocolos, como HTTP ou DNS.

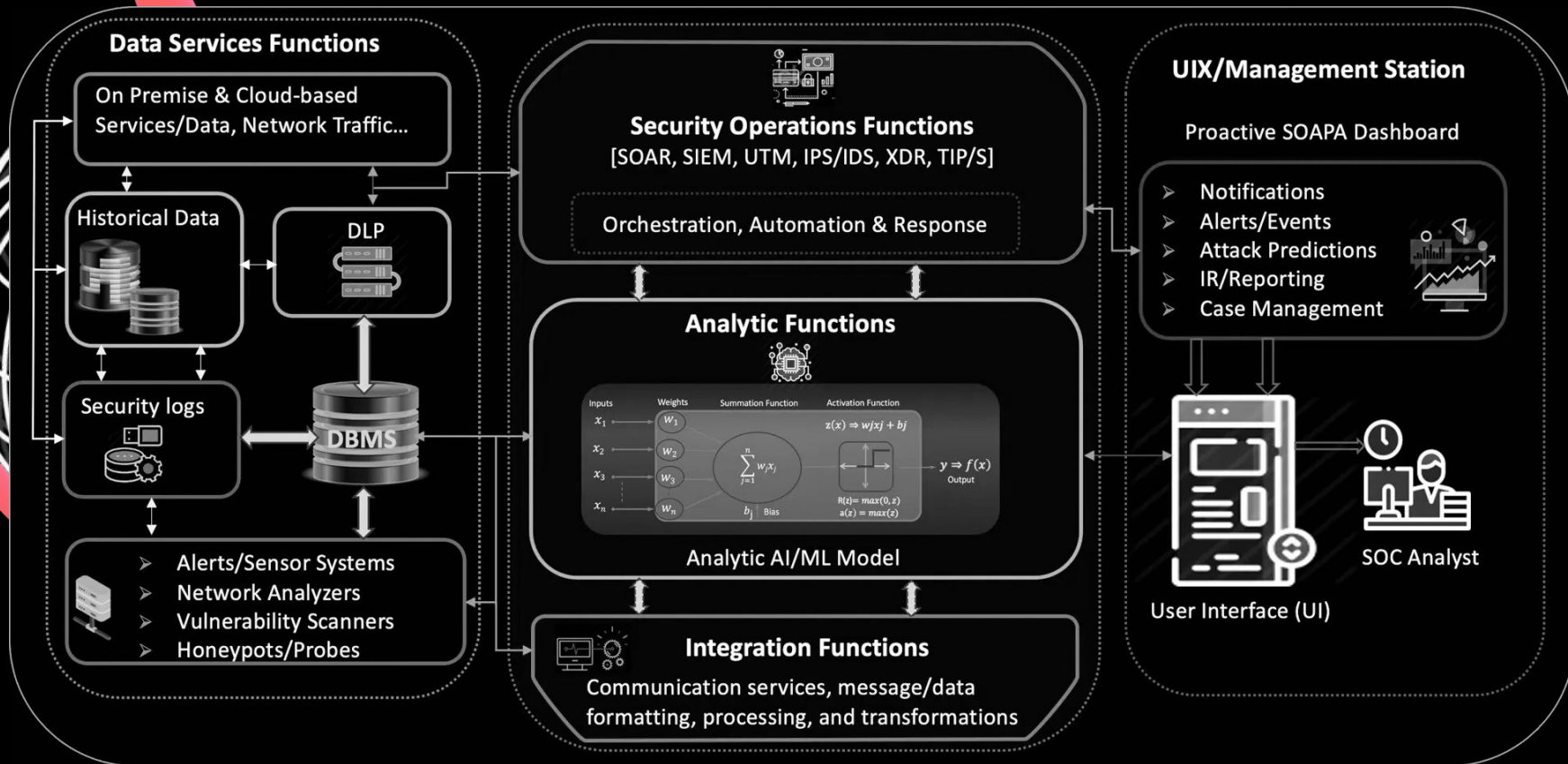
# Quais são os factores observáveis que reduzem a incerteza da interferência central do dispositivo comprometido?

## Comportamento de Rede Anómalos:

- Identificar hosts dentro de uma rede que estão a apresentar comportamento anómalo, como o envio de grandes quantidades de dados ou o acesso a recursos invulgares.



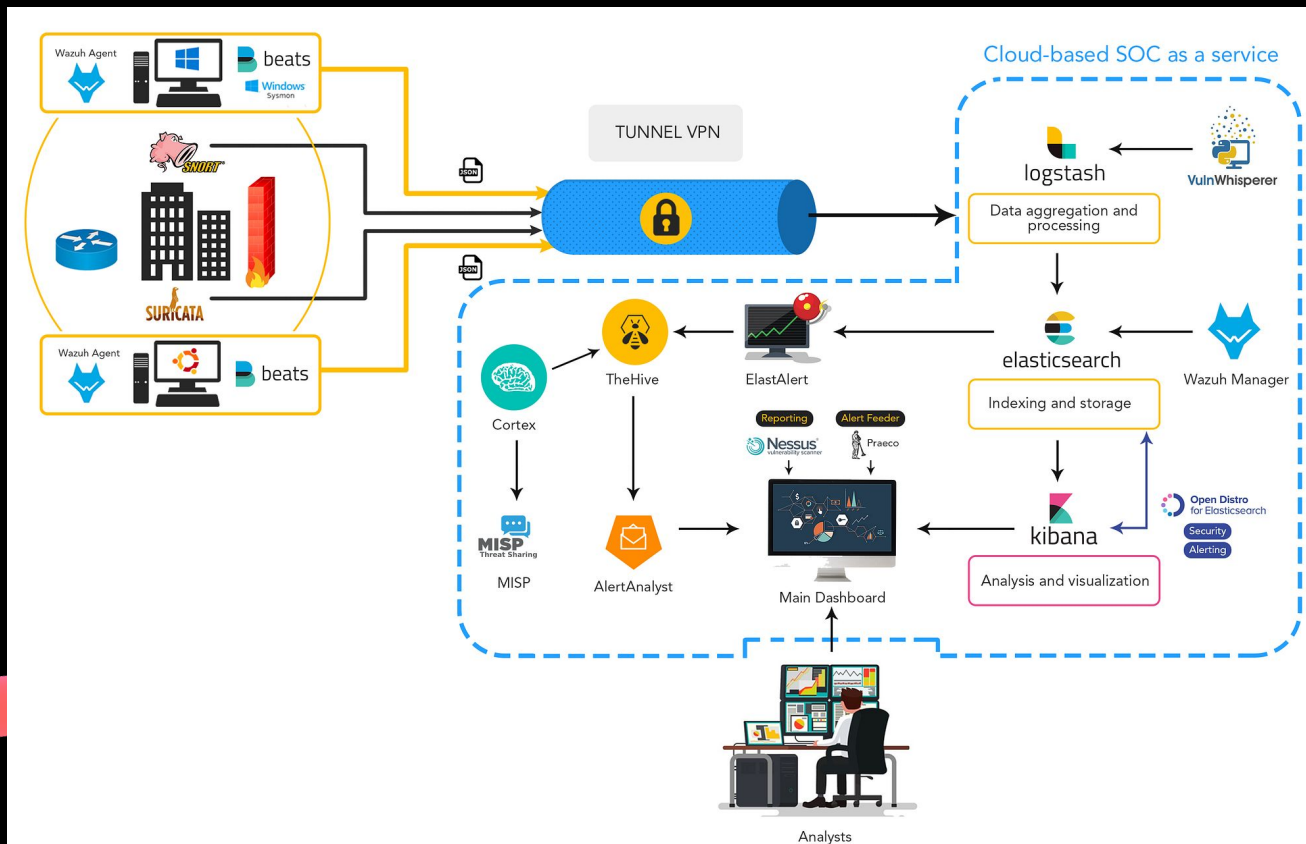
# Na prática o que tudo isso significa?



# Ainda sobre CTI

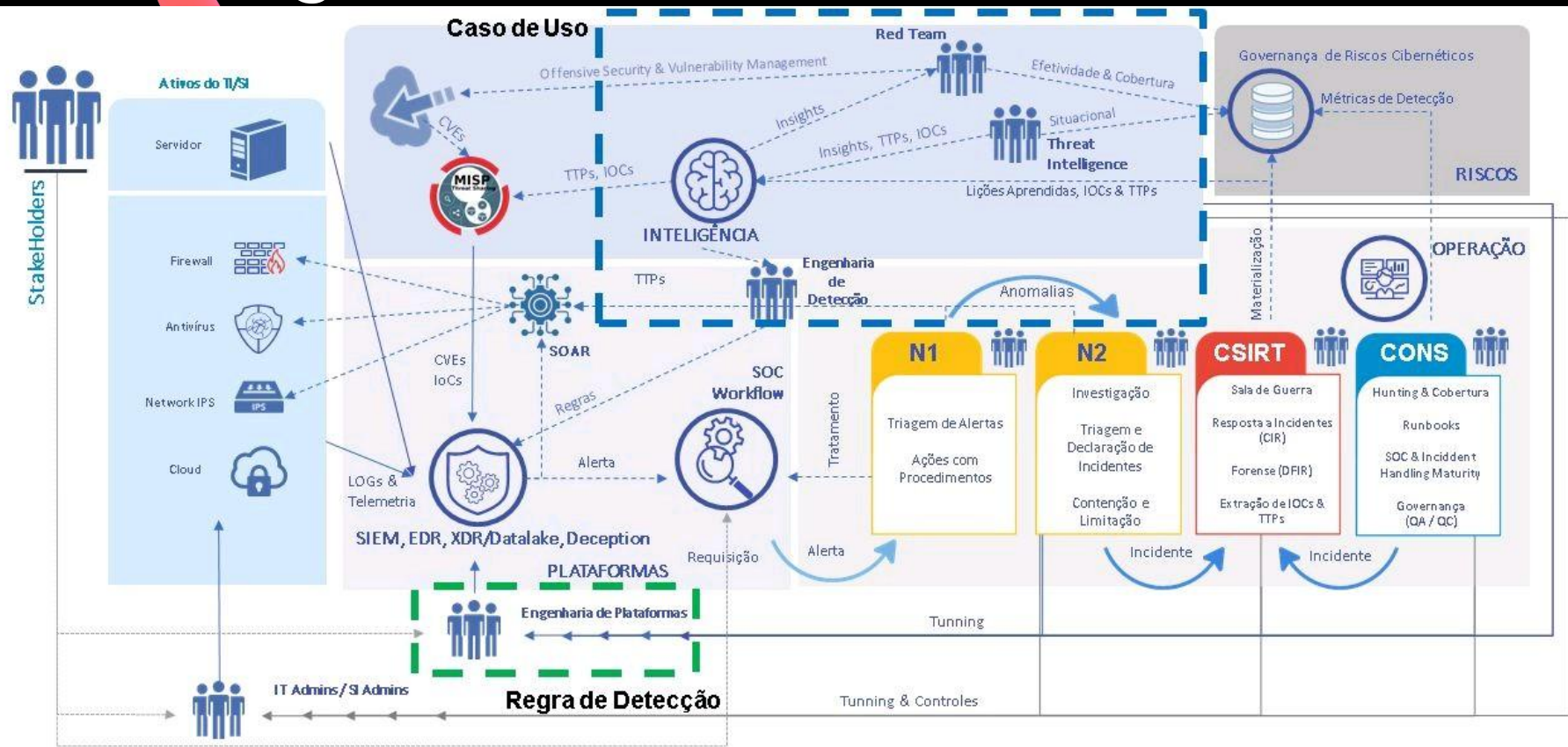


# Redução de processos manuais





# Na engenharia como funciona?



# Conclusion



# Q/A



<https://www.linkedin.com/in/arlindo0x73/>

