CIBERSEGURANÇA
NA ADMINISTRAÇÃO PÚBLICA

ENCONTROS SAMA 2020

26 ABRIL 14h30-17h15 LISBOA

ama AGÊNCIA PARA A MODERNIZAÇÃO ADMINISTRATIVA
COMPETE 2020
PORTUGAL 2020
UNIÃO EUROPEIA Fundo Europeu de Desenvolvimento Regional

Serviço de Gestão Alargada do Conhecimento Situacional e Operacional do Ciberespaço Nacional

ENCONTROS SAMA 2020

- **Desafio:**

  –

  Como identificar websites associados a Phishing..,

  … ANTES …

  de serem disseminadas campanhas (Phishing/Smishing) pelos Cidadãos

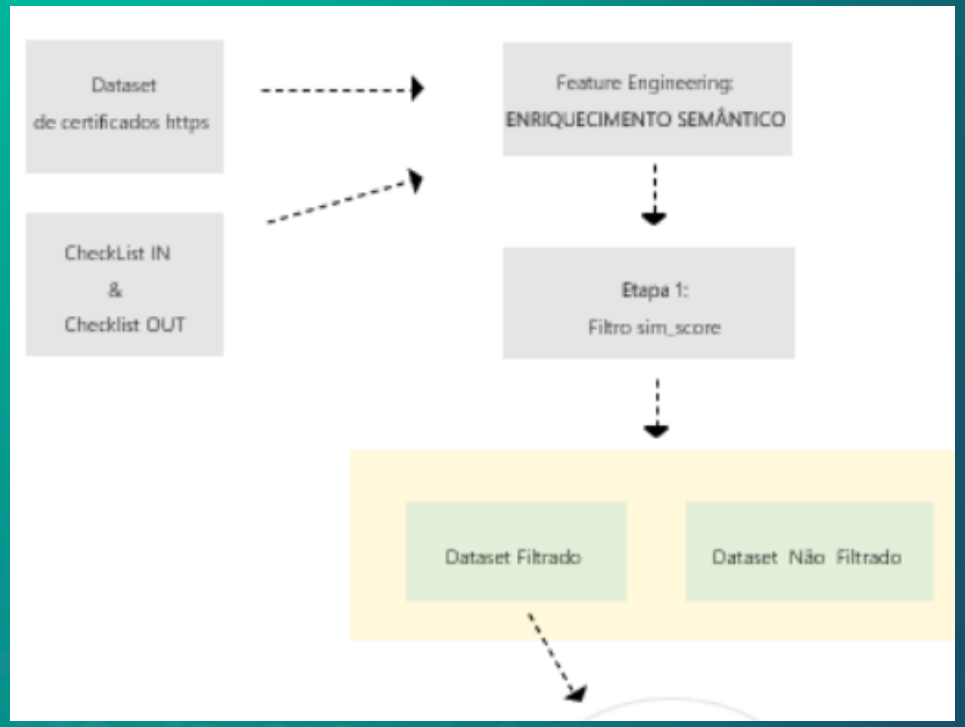
Serviço de Gestão Alargada do Conhecimento Situacional e Operacional do Ciberespaço Nacional

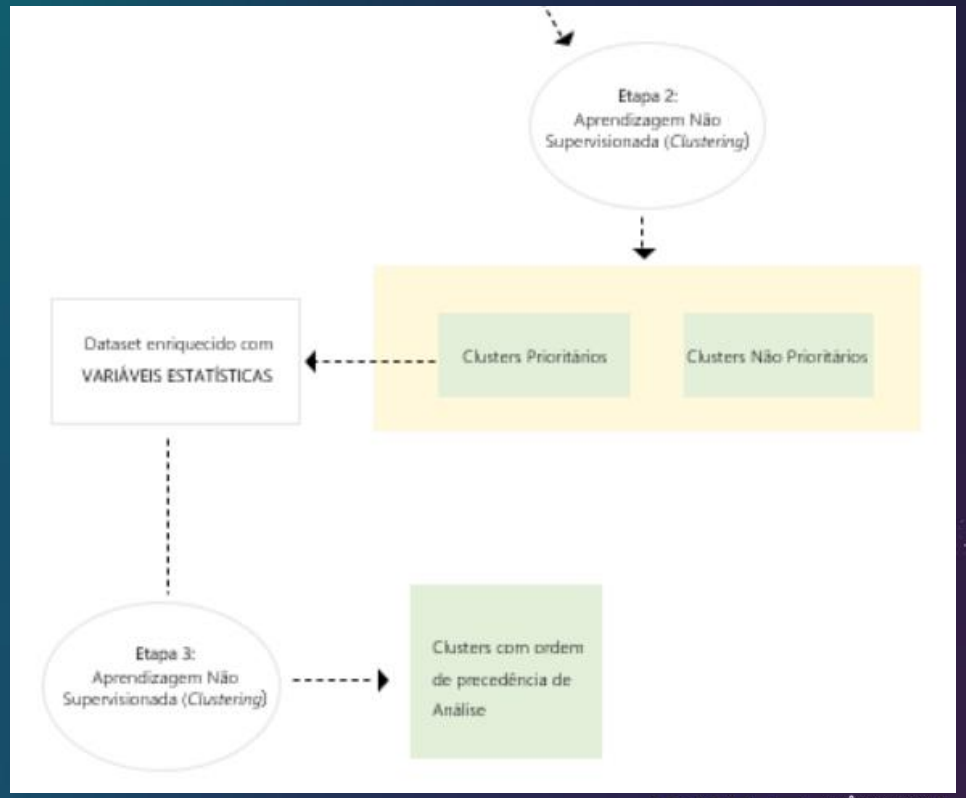
ENCONTROS SAMA 2020

- **Potencial Solução:**
  - 1- Sistema que verifica o registo de certificados
  - 2- Através de listas (semântica) dá uma classificação de domínios
  - 3- Definimos os *Clusters* prioritários
  - 4- Enriquecemos a informação com Estatística
  - 5- Analisamos os resultados
  - 6- Melhoramos o sistema com o *feedback* (aprendizagem supervisionada)

Serviço de Gestão Alargada do Conhecimento Situacional e Operacional do Ciberespaço Nacional

ENCONTROS SAMA 2020





ama AGÊNCIA PARA A MODERNIZAÇÃO ADMINISTRATIVA
COMPETE 2020
PORTUGAL 2020
UNIÃO EUROPEIA
Fundo Europeu de Desenvolvimento Regional

- **Fase 1 e 2:**
  - ***Streaming*** de uma fonte de informação de certificados

    Listas (semântica)
    - Expressões relevantes para o Ciberespaço Int. Nacional

    Aplicação de entropia *Shannon*
    - Aleatoriedade da palavra

Obtemos: um valor (*score*) para a semelhança com as palavras incluídas na checklist

- **Fase 3 e 4:**
  - Definidos os *clusters* **prioritários**

Enriquecemos a informação com:
- Rácio do número de vezes da *checklist* no domínio
- Estatística do *host* (presente em eventos maliciosos)
- Estatística relativa ao dono do domínio
- Estatística de e-mail utilizado pelo dono do domínio
- Configurações de DNS (DNS, MX, SPF, outros TXT)
- Estatística do *registar*
- Número de domínios por certificado
- CA utilizada para emissão dos certificados

- **Fase 5 e 6:**
  - **Novos *clusters* prioritários** – informação para validar

    Com base na fase anterior damos o *feedback* à aplicação de forma a que possa "**aprender**" e catalogar melhor (aperfeiçoamento do modelo)

- **Próximos passos:**
  - *Clusters* que não são prioritários

    Desafio do **volume de dados crescente**

    No enriquecimento de dados:
    - Algumas variáveis não retornam dados
    - Adicionar fontes externas de validação (fontes de Phishing, etc...)
    - Listas de validação atualizadas com base na clusterização e nos eventos diários

    Aperfeiçoar o modelo com o acervo de dados que o CERT.PT tem (desde 2014) observáveis a incidentes
  - Adicionar mais fontes de informação