

Process e Solution Design

Case Técnico RPA

Módulo A – Coleta Web & Normalização

Módulo B – Sistema de Newsletter

1. PDD – Process Definition Document

1.1. Visão Geral do Processo

Este documento descreve a automação do processo de **coleta, normalização e distribuição de notícias** oriundas do portal Yahoo News, com o objetivo de eliminar atividades manuais repetitivas realizadas pela área de Comunicação e Inteligência de Mercado.

O processo foi dividido em dois fluxos automatizados (Dispatcher e Performer), comunicando-se por meio de **fila no UiPath Orchestrator**, garantindo escalabilidade, rastreabilidade e padronização.

1.2. Processo AS-IS

Atualmente, o processo é executado manualmente por analistas, conforme descrito abaixo:

1. Acesso manual ao site <https://news.yahoo.com>.
2. Identificação visual das 5 notícias do bloco “Stories for you”.
3. Cópia manual de título, resumo e link.
4. Ajustes manuais de texto (remoção de caracteres inválidos).
5. Classificação subjetiva do tema da notícia.
6. Consolidação em planilhas locais.
7. Envio manual de e-mails para diferentes grupos.

Tempo médio diário: 4 a 6 horas

Principais erros identificados:

- Erros de digitação e formatação;
- Links incorretos;
- Dados duplicados;
- Classificação inconsistente de temas;
- Falta de histórico estruturado e rastreável.

1.3. Processo TO-BE

Com a automação, o processo passa a ser executado da seguinte forma:

- Um robô **Dispatcher** coleta e normaliza as notícias e publica os dados em fila.
- Um robô **Performer** consome os itens da fila e realiza a distribuição automática via e-mail, conforme o tema da notícia.

Tempo estimado após automação: ~30 minutos (processo não assistido).

1.4. Escopo da Automação

Incluído no escopo:

- Coleta automática de notícias no Yahoo News;
- Normalização e padronização dos dados;
- Geração de arquivo CSV diário;
- Publicação e consumo de fila no Orchestrator;
- Direcionamento automático de e-mails por tema;
- Utilização de configuração externa (Config.xlsx).

Fora do escopo:

- Criação ou manutenção de conteúdo editorial;
- Aprovação manual das notícias;
- Alterações visuais no portal de origem.

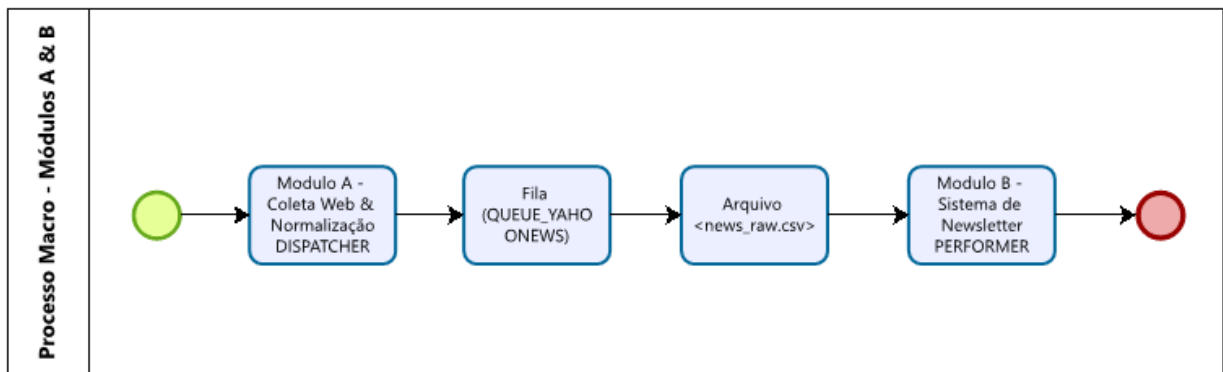
1.5. Benefícios Esperados

- Redução de tempo de até **85%**;
- Eliminação de erros manuais;
- Dados estruturados e reutilizáveis;
- Rastreabilidade completa via Orchestrator;
- Facilidade de manutenção e escalabilidade;
- Operação 24x7 sem intervenção humana.

2. SDD – Solution Design Document

2.1. Arquitetura da Solução

A solução foi desenvolvida utilizando UiPath REFramework, com separação clara entre Dispatcher e Performer, comunicando-se por meio de fila no UiPath Orchestrator.



2.2. Ferramentas e Componentes

- **RPA Tool:** UiPath Studio
- **Framework:** REFramework
- **Orquestração:** UiPath Orchestrator
- **Queue:** QUEUE_YAHOONEWS
- **Configuração:** Config.xlsx
- **Formato de saída:** CSV

2.3. Módulo A – Dispatcher (Coleta & Normalização)

Responsabilidades:

- Acessar <https://news.yahoo.com>
- Identificar o bloco “**Stories for you**”
- Extrair as 5 primeiras notícias
- Coletar os seguintes campos:

- titulo
 - resumo
 - tema
 - fonte
 - tempo_leitura
- Normalizar textos
- Gerar arquivo: news_raw_<DATA>.csv
- Publicar cada notícia como um **Queue Item** na fila QUEUE_YAHOONEWS

Configurações via Config.xlsx:

- URL do site
- Quantidade de notícias capturadas
- Caminho de saída do arquivo CSV;
- Nome da fila no Orchestrator;
- Grupos de emails
- Tema da notícia pré-definido

2.4. Módulo B – Performer (Sistema de Newsletter)

Responsabilidades:

- Consumir itens da QUEUE_YAHOONEWS
- Avaliar o campo **tema**
- Aplicar a regra de roteamento:
 - Tema = <tema definido pelo usuário> ex. *Politics* → Grupo de e-mail para o tema definido pelo usuário
 - Demais temas → Grupo de e-mail geral
- Recuperar listas de e-mail a partir de fonte externa configurável (não hardcoded)
- Montar e enviar e-mails automaticamente
- Atualizar status do item da fila (Successful / Failed)

2.5. Tratamento de Exceções

- **Business Exceptions:**
 - Tema não identificado
 - Lista de e-mails vazia ou inexistente
- **System Exceptions:**
 - Falha de acesso ao site
 - Erro de carregamento de página
 - Falha de comunicação com Orchestrator
- Todas as exceções são tratadas conforme padrão do REFramework, com logging e retry configurável

2.6. Segurança e Governança

- Credenciais e parâmetros sensíveis armazenados como **Assets no Orchestrator**
- Logs centralizados no Orchestrator
- Controle de acesso por roles
- Possibilidade de auditoria completa por execução e item de fila.

3. Conclusão

A automação proposta substitui um processo manual crítico por uma solução robusta, escalável e alinhada às boas práticas de RPA corporativo. O uso do REFramework, aliado à comunicação via Queue no Orchestrator, garante confiabilidade, manutenção simplificada e expansão futura do processo.