



# **Relatório do Trabalho Prático 1**

Integração de Sistemas Informáticos

João Arantes

a21600

a21600@alunos.ipca.pt

**Licenciatura em Engenharia Sistemas Informáticos**

**3ºano**

Barcelos | Outubro, 2024

## Índice de figuras

Figura 1 - ExtratMeteorologia .....	9
Figura 2 - Html .....	10
Figura 3 - Get_Meteorologia.....	11
Figura 4 - Job.....	12
Figura 5 - QR code.....	14

# Índice

Índice.....	2
Introdução.....	4
Enquadramento.....	5
Problema.....	6
Estratégia utilizada.....	7
Transformações.....	9
Transformação ExtratMeteorologia.....	9
Transformação Html.....	10
Job Get_Meteorologia.....	11
Job Automatizado.....	12
Explicação .....	13
Vídeo com demonstração (QR Code).....	14
Conclusão.....	15

# Introdução

Este projeto de Integração de Sistemas Informáticos centra-se na criação de um processo de ETL (Extração, Transformação e Carregamento) aplicado à meteorologia, permitindo a recolha, transformação e armazenamento de dados climáticos de forma automatizada e eficiente. Utilizando o Pentaho Kettle, o projeto realiza a extração de dados meteorológicos a partir da API da Open Weather Map, assegurando que as informações tratadas estejam sempre atualizadas e relevantes. O processo inclui a importação e exportação de dados em formatos XML, TXT e Excel, facilitando a interoperabilidade entre sistemas.

Além disso, o desenvolvimento de um Job com acesso a serviços de e-mail possibilita a comunicação eficaz na partilha de informações meteorológicas relevantes. A estrutura do projeto assegura a consistência e a qualidade dos dados tratados, com a implementação de transformações que incluem a limpeza e normalização de caracteres, bem como a ordenação dos dados. Ao explorar o acesso a APIs remotas, este projeto oferece uma solução prática e inovadora para o tratamento de grandes volumes de dados meteorológicos.

# Enquadramento

A meteorologia desempenha um papel crucial na sociedade contemporânea, influenciando decisões em diversos setores, desde a agricultura até a gestão de recursos hídricos e o planeamento urbano. Com o aumento da disponibilidade de dados climáticos em tempo real, a capacidade de processar, analisar e interpretar essas informações tornou-se fundamental para a tomada de decisões informadas. Este projeto surge no contexto da crescente necessidade de integrar e automatizar o tratamento de dados meteorológicos, utilizando tecnologias modernas de integração de sistemas.

O uso do Pentaho Kettle como ferramenta de ETL permite a extração eficiente de dados de várias fontes e a transformação necessária para garantir a qualidade e a relevância dos dados. A capacidade de serializar dados em formatos amplamente utilizados, como XML e TXT, assegura a interoperabilidade entre diferentes sistemas e aplicações. Além disso, a implementação de um Job que permite o envio automatizado de alertas e informações relevantes por e-mail garante que as partes interessadas recebam dados críticos de forma oportuna.

A exploração de APIs remotas oferece a oportunidade de enriquecer o projeto com dados adicionais, ampliando a base de conhecimento disponível. Neste sentido, este projeto não é apenas uma solução técnica, mas uma resposta a um desafio real na área da meteorologia, contribuindo para uma melhor compreensão e utilização dos dados climáticos disponíveis.

# Problema

A crescente quantidade de dados meteorológicos disponíveis em tempo real apresenta um desafio significativo para a sua manipulação e análise eficaz. Muitas organizações e indivíduos têm dificuldade em integrar essas informações de forma a obter insights valiosos e tomar decisões informadas. O problema principal reside na falta de uma solução automatizada que permita a extração, transformação e armazenamento de dados meteorológicos de diversas fontes, garantindo a sua qualidade e acessibilidade.

Além disso, a diversidade de formatos de dados (como XML e TXT) pode complicar a interoperabilidade entre sistemas, dificultando a integração de diferentes fontes de dados. A necessidade de um processo eficiente de serialização e deserialização é vital para assegurar que os dados possam ser utilizados de forma consistente em diferentes aplicações.

Por último, a comunicação das informações meteorológicas relevantes é frequentemente limitada a métodos manuais, resultando em atrasos e possíveis erros na partilha de dados críticos. A falta de um sistema que permita o envio automatizado de alertas e informações climáticas através de serviços de e-mail contribui para a ineficiência na utilização destes dados.

Este projeto visa resolver esses problemas, proporcionando uma solução integrada que não apenas automatiza o tratamento de dados meteorológicos, mas também melhora a sua acessibilidade e utilização, permitindo uma resposta mais rápida e informada às mudanças climáticas.

## Estratégia utilizada

Para abordar os desafios identificados, o projeto adota uma estratégia estruturada que envolve a implementação de um processo de ETL (Extração, Transformação e Carregamento) utilizando a ferramenta Pentaho Kettle. Esta estratégia contempla os seguintes operadores e processos:

1. **Extração de Dados:** A fase inicial do processo envolve a utilização de um cliente REST para a extração de dados meteorológicos da API da Open Weather Map. Esta abordagem permite obter dados em tempo real, assegurando que as informações tratadas estejam sempre atualizadas e relevantes. O passo inicial é realizado através da operação Generate Rows, que se liga à API para gerar as linhas de dados.
2. **Transformação de Dados:** Após a extração, os dados passam por uma série de transformações para garantir a sua consistência e qualidade. Os passos envolvidos incluem:
  - JSON Input para ler os campos relevantes (id, main, description, e date) dos dados extraídos.
  - Modified JavaScript Value para limpeza e normalização de caracteres, assegurando que os dados estejam em um formato adequado para análise.
  - Select Values para selecionar e renomear os campos conforme necessário.
  - Sort Rows para organizar os dados de acordo com critérios específicos, facilitando a análise e interpretação.
  - Os dados transformados são, então, exportados para formatos XML, TXT e Excel, permitindo sua utilização em diversas aplicações.

3. **Processamento de HTML:** O projeto inclui uma transformação específica para lidar com ficheiros HTML. Esta transformação começa com um TXT Input que lê o conteúdo de um ficheiro HTML, seguido por um Group By e um SetString para definir a variável `ambient`, permitindo a extração e análise de informações relevantes contidas no HTML.
4. **Envio de Alertas:** A comunicação das informações meteorológicas é realizada através de um Job específico, `Get_Meteorologia`, que integra várias operações:
  - Inicia com o passo `Delete File` para remover qualquer ficheiro HTML obsoleto.
  - Executa a `Transformation ExtratMeteorologia` para assegurar que os dados meteorológicos estão atualizados.
  - Verifica a existência do ficheiro `meteorologia.xml` utilizando o passo `File Exists`.
  - Aplica uma `XSL Transformation` para formatar os dados conforme necessário.
  - Realiza uma segunda transformação HTML para preparar os dados para envio.
  - Por fim, envia as informações meteorológicas por e-mail, garantindo uma comunicação eficiente com as partes interessadas.
5. **Automatização do Processo:** Para simplificar a execução de todas as etapas, um Job adicional automatiza o processo geral, iniciando com a execução da `Transformation ExtratMeteorologia`, seguida pela execução do Job `Get_Meteorologia`. Esta automação assegura que todo o processo de ETL seja realizado de forma contínua e eficiente.

A combinação destes operadores e processos cria uma solução integrada e automatizada que simplifica o tratamento de dados meteorológicos, melhora a eficiência na sua utilização e comunicação, e proporciona uma resposta mais rápida às mudanças climáticas.



# Transformações

Nesta seção, são apresentadas as transformações utilizadas no projeto, com uma descrição detalhada de cada uma. As transformações são fundamentais para o processo de ETL, garantindo que os dados meteorológicos sejam extraídos, tratados e disponibilizados de forma eficaz.

## Transformação ExtratMeteorologia

**Descrição:** Esta transformação é responsável por extrair os dados meteorológicos da API da Open Weather Map, realizar a limpeza e normalização dos dados, e exportá-los em diversos formatos.

**Passos:**

1. **Generate Rows:** Gera as linhas de dados com base nas solicitações à API.
2. **REST Client:** Realiza a chamada à API para obter os dados meteorológicos.
3. **JSON Input:** Lê os campos id, main, description e date dos dados retornados.
4. **Modified JavaScript Value:** Limpa e normaliza caracteres nos dados para garantir sua integridade.
5. **JSON Output:** Formata os dados para saída em formato JSON.
6. **Select Values:** Seleciona e renomeia campos conforme necessário.
7. **Sort Rows:** Ordena os dados de acordo com critérios definidos.
8. **Exportação:** Os dados são extraídos para os formatos XML, TXT e Excel.

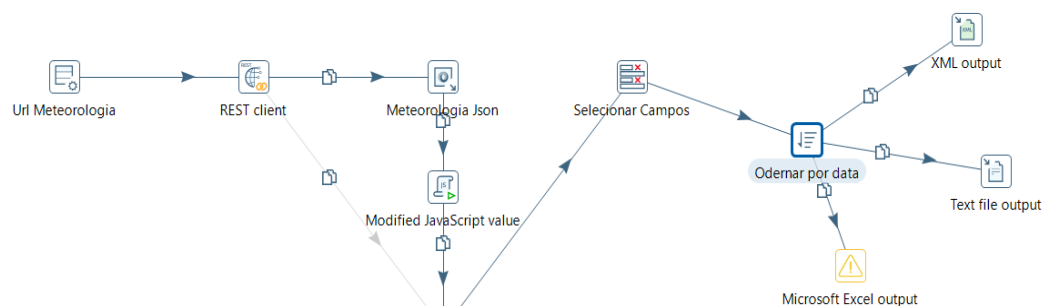


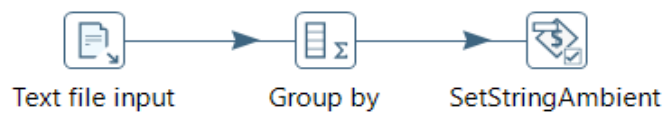
Figura 1 - ExtratMeteorologia

## Transformação Html

**Descrição:** Esta transformação lida com a extração e tratamento de informações a partir de um ficheiro HTML.

**Passos:**

1. **TXT Input:** Lê o conteúdo do ficheiro HTML.
2. **Group By:** Agrupa os dados conforme critérios definidos.
3. **SetString:** Define a variável ambient com os dados agrupados.



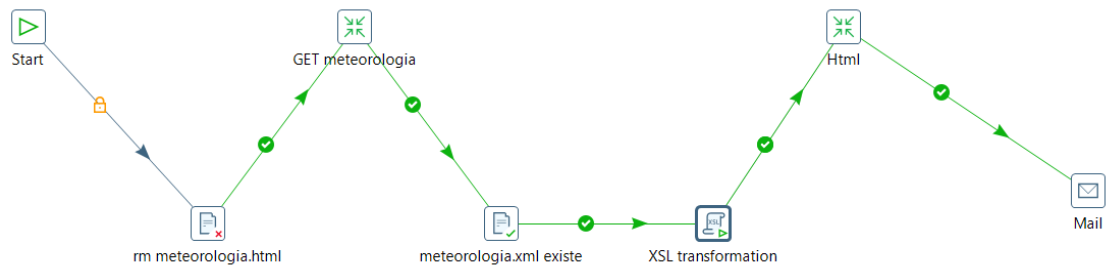
*Figura 2 - Html*

## Job Get\_Meteorologia

**Descrição:** Este Job automatiza a execução de várias operações para garantir que os dados meteorológicos sejam atualizados e enviados por e-mail.

**Passos:**

1. **Start:** Início do Job.
2. **Delete File:** Remove o ficheiro HTML obsoleto.
3. **Transformation ExtratMeteorologia:** Executa a transformação para extrair os dados meteorológicos.
4. **File Exists:** Verifica a existência do ficheiro meteorologia.xml.
5. **XSL Transformation:** Aplica uma transformação XSL aos dados.
6. **Transformation Html:** Prepara os dados para envio.
7. **Mail:** Envia as informações meteorológicas por e-mail.



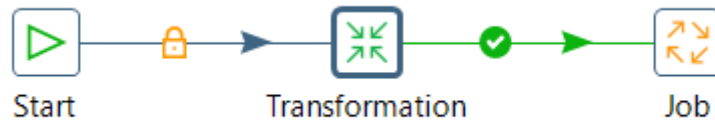
*Figura 3 - Get\_Meteorologia*

## Job Automatizado

**Descrição:** Este Job automatiza o processo de ETL, integrando a execução da transformação e do Job de envio.

**Passos:**

1. **Start:** Início do Job.
2. **Transformation ExtratMeteorologia:** Executa a transformação para extrair os dados meteorológicos.
3. **Job Get\_Meteorologia:** Executa o Job de envio de dados.



*Figura 4 - Job*

## Explicação

As transformações implementadas no projeto são essenciais para garantir a correta manipulação dos dados meteorológicos. A Transformação ExtratMeteorologia assegura que os dados sejam extraídos de forma eficiente da API, tratados para garantir a sua qualidade, e disponibilizados em diferentes formatos que facilitam a sua utilização em várias aplicações.

A Transformação Html complementa o processo, permitindo a extração de informações de um ficheiro HTML e a preparação de dados relevantes para posterior análise.

O Job Get\_Meteorologia integra várias operações, garantindo que os dados sejam atualizados, formatados e enviados por e-mail de forma automatizada. Finalmente, o Job Automatizado proporciona uma visão geral do processo de ETL, permitindo uma execução contínua e eficiente das operações necessárias para o tratamento dos dados meteorológicos.

Vídeo com demonstração (QR Code)



*Figura 5 - QR code*

<https://drive.google.com/file/d/12zwxj58x4PltXCnXEPoQjKCmZOW1T11L/view>

## Conclusão

O projeto de integração de sistemas informáticos sobre meteorologia representa um esforço significativo na extração, tratamento e disponibilização de dados meteorológicos em tempo real. Através da utilização da API da Open Weather Map, conseguimos garantir a obtenção de informações atualizadas e relevantes, que são essenciais para a análise e a tomada de decisões em diversas áreas.

A implementação de transformações eficientes, como a `ExtratMeteorologia`, possibilitou não apenas a coleta de dados, mas também a sua limpeza e normalização, assegurando que a informação esteja em condições ideais para uso. Adicionalmente, o tratamento de dados extraídos de um ficheiro HTML enriqueceu o nosso projeto, proporcionando uma visão mais ampla sobre as variáveis ambientais.

A automação dos processos através dos Jobs desenvolvidos, especialmente o `Get_Meteorologia`, facilitou a atualização regular e a disponibilização dos dados, permitindo o envio automático de informações críticas por e-mail. Esta abordagem não só otimiza o tempo de resposta a novas informações, mas também assegura a continuidade na gestão dos dados meteorológicos.

Por fim, a inclusão de um vídeo de demonstração serve como uma ferramenta visual que complementa a documentação, permitindo que os interessados compreendam facilmente o funcionamento do sistema e as suas potencialidades. Este projeto não apenas demonstra a viabilidade técnica da integração de sistemas informáticos, mas também a sua aplicabilidade prática em contextos do mundo real.

Acreditamos que as soluções apresentadas neste projeto têm o potencial de ser expandidas e aprimoradas, podendo ser adaptadas para diferentes cenários que exijam a coleta e análise de dados de forma automatizada e eficiente.-