**Uma imagem com texto, verde, captura de ecrã, Tipo de letra

Os conteúdos gerados por IA podem estar incorretos.**

**Planeamento – Apoio à decisão**

José Carlos Lima Gonçalves

*Professor*

Luis Ferreira

Projeto apresentado ao Instituto Politécnico do Cávado e do Ave, no âmbito da disciplina de Integração de Sistemas de Informção do curso de Licenciatura de Engenharia de Sistemas informáticos

Outubro, 2025

Enquadramento

Num contexto de crescente digitalização e complexidade nos modelos e processo de negócio, torna-se essencial garantir que os dados são não apenas recolhidos, mas também tratados e disponibilizados de forma eficiente. Este projeto insere-se nesse desafio, explorando três áreas essenciais.

ETL (Extract, Transform, Load), o processo de extração, transformação e carregamento de dados é central para garantir que a informação proveniente de diversas fontes é normalizada e preparada para consumo por sistemas secundários.

A utilização de ferramentas como o Pentaho data Integration, Knime, Power Query, Power Automator, são algumas ferramentas que permitem facilitar e automatizar esses processos.

A segunda área é a Arquitetura Modular que promovem a manutenção, escalabilidade, ambiente de testes e integração. Cada componente do processo pode ser isolado e modular de forma a reduzir a complexidade e aumentar a fiabilidade.

E por fim, a Interoperabilidade, ou seja, a capacidade de diferentes sistemas comunicarem entre si é crucial para evitar dependências tecnológicas. Este projeto foca-se em criar formatos de dados que podem ser facilmente integrados por sistemas externos promovendo a flexibilidade e a abertura tecnológica, no entanto apenas foram abordadas uma ínfima parte da potencialidade das ferramentas acima mencionadas.

O trabalho surge no decorrer da disciplina de Integração de Sistemas de Informação do curso de Licenciatura em Engenharia de Sistemas Informáticos a decorrer em Barcelos no Instituto Politécnico do Cávado e do Ave a ser lecionado em ambiente pós-laboral.

Context

In a landscape of increasing digitalization and business process complexity, it is essential not only to collect data but also to process and deliver it efficiently. This project addresses that challenge by focusing on three key areas.

ETL (Extract, Transform, Load), the ETL process is central to ensure that data from various sources is normalized and prepared for consumption by secondary systems.

Tools such as Pentaho Data Integration, KNIME, Power Query and Power Automate are among the solutions that help streamline and automate these workflows.

Modular Architecture, this approach promotes maintainability, scalability, testing environments and integration. Each component of the process can be isolated and modularized to reduce complexity to enhance reliability.

And last Interoperability, that is the ability of different systems to communicate with one another is crucial to avoiding technological lock-in. This project focuses on creating data formats that can be easily integrated by external systems, promoting flexibility and technological openness. However, only a small portion of the potential of the tools has been explored.

This project was developed as part of the degree “Engenharia de Sistemas Informáticos” in the subject of information systems integration at IPCA in Barcelos.

Projecto

Pretende-se com estre projeto complementar um segundo sistema de automação já existente que utiliza o Power Automate que diariamente envia os requisitos necessários de produção medidos por número de pessoas, por intervalos de 15 minutos e um segundo email que com base na informação existente no sistema de planeamento envia os horários do dia de cada pessoa para cada linha de negócio.

O processo de Power Automate pega nos anexos correspondentes e coloca-os numa pasta pré-definida.

Os ficheiros gerados são divididos em dois tipos, excel e csv, o que se pretende é pegar nas duas informações convergir e medir os riscos de falta de cobertura automatizando o processo, pretende-se também gerar um email com os intervalos com risco de falta de produção guardando uma cópia com os dados combinados na sua totalidade para histórico e futura integração numa base de dados para análise de tendências, e um segundo com os dados filtrados apenas pelos intervalos que estão a descoberto.

Ferramentas utilizadas

Ao longo do desenvolvimento do projeto houve alguma indecisão quando à ferramenta devido à curva de aprendizagem e facilidade de utilização, esta acabou por pender para o KNIME versão 5.7.0 pela facilidade e menor curva de aprendizagem naquelas que seriam as tarefas mais simples, no entanto a gestao de potenciais erros (log e saídas de processo) tornou-se algo complexo e a continuar o seu desenvolvimento, questão que no Pentaho veio a tornar-se mais simples nessas questões.

KNIME – Planeamento

*Ficheiros de input:*

*Requirements.csv*  
Campos de dados  
Date: data a que se referem os dados

Interval: início de cada intervalo de 15 min

Line of Business: Linha de negocio

Requirements: número de pessoas produtivas necessárias

*Horários – um ficheiro por cada linha de negócio em formato xlsx (Excel)*  
Campos de dados  
Date: data a que se referem os dados

IEX ID: Identificador de colaborador

Name: Identificador de colaborador

Line of Business: Linha de negócio

Shift Start: Inicio de turno

Shift End: Fim de turno

Visão do projeto por modulos

Uma imagem com texto, diagrama, file, captura de ecrã

Os conteúdos gerados por IA podem estar incorretos.

Figura 1

*Requirements.csv*

1. Identificar localização onde o ficheiro estará disponível
2. Preparar os dados para analise futura para os mesmos tipos de linhas de negócio para diferentes línguas categorizando cada linha de negócio na sua língua e nível de serviço (Lob short).
3. Parsing dos campos (IEX ID – number, Data – datetime, Interval- Time.
4. Adicionar o fim do intervalo para futuro mapeamento com a hora de entrada e saída dos colaboradores para contagem acumulada. (adicionar 15 minutos à hora de início de intervalo, conforme requisito)
5. Criar o modulo de forma a isolar o processo conforme imagem abaixo

Uma imagem com diagrama, file, texto, captura de ecrã

Os conteúdos gerados por IA podem estar incorretos.

Figura 2

*Horarios.xlsx*

1. Identificar localização onde o ficheiro estará disponível, a utilização do componente Excel Reader inclui o parsing aquando da leitura dos ficheiros
2. Preparar os dados para analise futura para os mesmos tipos de linhas de negócio para diferentes línguas
3. Ordenação dos dados para contagem acumulada
4. Divisão da informação em dois grupos para conseguir contagem de entradas e saídas por intervalo por linha de negócio
5. Para cada um dos processos criou-se um loop para fazer a contagem acumulada de forma a zerar para cada uma das linhas de negócio.
6. Renomeação das colunas calculadas
7. Criar o modulo de forma a isolar o processo conforme imagem abaixo

Uma imagem com diagrama, file, texto, captura de ecrã

Os conteúdos gerados por IA podem estar incorretos.

Figura 3

*Operação de Join e Cálculo*

1. Full Right Join da tabela de Requirements com Horarios de saídas
2. Substituição dos campos vazios nos horários de saída por zero (acumulado)
3. Segundo Full Right Join do resultado do processo anterior com as entradas acumuladas
4. Substituição dos campos vazios nos horários de entrada por zero (acumulado)
5. Ordenação por linha de negócio e intervalo
6. Criação de um loop de processo de repetição para soma acumulada para entradas e saídas
7. Cálculo da diferença entre os requisitos, entradas e saídas de forma a gerar a cobertura vs requisitos por intervalo
8. Criar o modulo de forma a isolar o processo conforme imagem abaixo

Uma imagem com captura de ecrã, diagrama, file

Os conteúdos gerados por IA podem estar incorretos.

*Operações de Output (Json, excel e email)*

1. Filtrar o resultado da operação anterior em que o “over-under” seria abaixo de zero, ou seja, os requisitos eram superiores à cobertura existente.
2. Conversão da tabela em Json, para integração com sistemas futuros para dashboard nos intervalos com baixa cobertura
3. Escrita do resultado num ficheiro de excel guardado como temporário para envio por email, sempre que é enviado este é sobreposto.
4. Envio de email com anexo do ficheiro de excel gerado no processo anterior
5. Salvaguarda de toda a informação antes de ser filtrada num ficheiro de Excel com a data em que o processo corre para efeitos de histórico, para isto recorreu-se à data de requisitos existente no ficheiro de CSV de forma a manter coesa o nome do ficheiro contra a informação presente, aqui falta no entanto complementar com confirmação que só existe uma so data no ficheiro (por exemplo) uma vez que ira pegar sempre pegar no primeiro registo do ficheiro.

Uma imagem com texto, diagrama, file, captura de ecrã

Os conteúdos gerados por IA podem estar incorretos.

*Configuração de email*

1. Para efeitos de teste foi criada uma conta gmail ([zecvvtesteisi@gmail.com](mailto:zecvvtesteisi@gmail.com))
2. Configurada password de aplicação para ser utilizada de forma segura no knime
3. Configuração no modulo de Send Email, com o email, pw de aplicação e configuração da porta SSL
4. Configurada mensagem de corpo do email e anexo a ser enviado, a mensagem de corpo pode ser configurada dinamicamente utilizando variáveis de ambiente na execução do projeto.
5. Para o envio do anexo foi necessário recorrer a variáveis de ambiente para serem convertidas no caminho e nome do ficheiro a serem utilizados como referencia.

Uma imagem com texto, captura de ecrã, número, software

Os conteúdos gerados por IA podem estar incorretos.

Conclusão

KNIME foi uma bela surpresa pela simplicidade e os diversos módulos existentes mesmo ao nível de dasboarding, tendo explorando alguns, no entanto não conseguindo resolver todos os obstáculos que ocorreram pelo meio assim como algumas dificuldades com as variáveis de ambiente e a forma como elas comunicavam entre os diferentes módulos afetando o resultado final.

Trabalho futuro a ser necessário a ser desenvolvido mais centrado na gestao de erros e de registo de ocorrências ao longo do processo.

O opção pelo KNIME e a mudança para este teve a ver com o número de variáveis e curva de aprendizagem do pentaho o que levou ao atraso no inicio do projeto, sendo que o KNIME acelerou o desenvolvimento.  
Com base na experiencia profissional dependendo das ferramentas a usar, Power Query na minha opinião é mais simples que ambas as soluções para o utilizador final caso o projeto não seja de muita envergadura, no entanto é bastante mais limitada no que diz respeito a formatos e tecnologias envolvidas.

Com isto duas ferramentas com um potencial enorme especialmente neste âmbito em que diferentes sistemas libertam diferentes formatos de ficheiros que se visam a complementar para ajudar à toma de decisão.