

# Projeto do Módulo de Interação com o MW (Data Mining)



## Membros:

- Marcos Felipe
- Adonay Veigar

## 1. Descrição

Esse subsistema é responsável pela extração de dados do Matricula Web que serão enviados ao Banco de Dados. Ele é baseado em algoritmos de parserização, que quebra o código HTML da página em Tabelas, que podem ser utilizadas para adquirir as informações necessárias.

Sua implementação pode ser dividida em módulos:

- Extrair cursos: Extrai quais são os cursos disponíveis na UnB, assim como o número de créditos necessários para completar o curso.
- Extrair currículo do curso: Adquire informações relativas a quais são as matérias (Optativas e obrigatórias) dos cursos obtidos anteriormente, assim como o pré-requisito de cada disciplina, e o número de créditos obtidos ao cursá-la.
- Extrair oferta por departamento: Responsável por adquirir as informações a respeito de quais disciplinas estarão sendo ofertadas em cada departamento no respectivo semestre. Traz informações como código da disciplina, nome, turmas, horários e professores.
- Extrair histórico do aluno: Carrega o histórico do aluno (Disciplinas cursadas e respectivas menções, disciplinas obrigatórias a serem cursadas, número de créditos já feitos).
- Enviar ao BD: Envia as informações obtidas ao banco de dados.

## 2. Entradas e Saídas

- Código HTML das páginas no Matricula Web.
- Informações de como as tabelas do banco de dados estão dispostas, assim como o nome de cada uma, e seus respectivos atributos.



### 3. Ferramentas a serem utilizadas

- JSoup: Biblioteca Java responsável pela aquisição e quebra do código HTML das páginas do MW em tabelas.
- Algoritmos de tratamentos dos dados recebidos, criados pela equipe.

## Projeto do Módulo de Interação com o MW (Data Mining)



Versão	Modificações	Revisor
1.0	Criação do Documento	Marcos Felipe