|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Prioridade | Product Backlog | User Story # | User Story | Estimate (Hours) |
| 1 | Leitura de ficheiros |  | Faz a leitura de ficheiros. |  |
| 1 | Remover dígitos e caracteres de pontuação |  | Em cada ficheiro, remove toda a pontuação e dígitos numéricos. |  |
| 2 | Criar de uma matriz M |  | Cria uma matriz M que representa o número de ocorrências de cada palavra na coleção dos documentos. As linhas corresponderão aos documentos e as colunas corresponderão às palavras nos documentos. |  |
| 2 | Criar de uma matriz Q |  | Cria uma matriz equivalente a um array que será preenchido com a frase introduzida pelo utilizador. |  |
| 3 | Substituir dos valores da matriz M através do cálculo dado |  | Esta função irá substituir cada valor do elemento Mij pelo resultado do cálculo Mij= Mij log10(N/Np).  N – nº total de documentos;  Np- nº de documentos que contêm a palavra p (valor da coluna j da matriz). |  |
| 3 | Substituir dos valores da matriz Q através do cálculo dado |  | Repetir o processo acima descrito, mas aplicando agora à matriz Q. |  |
| 4 | Calcular o grau de semelhança |  | Através da expressão GrauSim=[ (somatório)(MLi\*Qi)]/[(raíz)(Mi)\*(raíz)(Qi^2).  M- array correspondente à linha I da matriz M;  Q- array correspondente à query de pesquisa. |  |
| 5 | Apresentar os resultados |  | Criação do ranking dos documentos para de seguida apresentar uma lista com todos os documentos analisados por ordem de similaridade. Também apresentará uma lista com todos os ficheiros com o grau de similaridade superior ao definido pelo utilizador e ainda uma lista com um limite máximo de ficheiros definido por input externo. |  |