

SISTEMA MULTI-AGENTE DE PESQUISA INFORMAÇÃO INTERNET

- **OBJECTIVO:** Auxiliar a pesquisa de informação de interesse na Internet
- **Retorna informação:**
 - A pedido do utilizador: Através da indicação de palavras-chave;
 - Através da análise da informação relevante das páginas que o utilizador está a navegar.

SMAPII

Sistema Multi-Agente de Pesquisa de Informação na Internet (SMAPII)

Página: Texto a pesquisar: "Information Retrieval" N.º Pág. 50

[d:\Tomcat\webapps\SMA\Paginas\Information_retrieval\Information_Retrieval_Research_-_SearchTools_Topics.htm](http://localhost:8080/SMA/Paginas/Information_retrieval/Information_Retrieval_Research_-_SearchTools_Topics.htm) 81 TFIDF: engine trec excellent berkeley

[d:\Tomcat\webapps\SMA\Paginas\Information_retrieval/pers-s-eng.htm](http://localhost:8080/SMA/Paginas/Information_retrieval/pers-s-eng.htm) 32 TFIDF: slovenian personalised gathered big

[d:\Tomcat\webapps\SMA\Paginas\Information_retrieval/music-ir_org_virtual_home_of_music_information_retrieval_research.htm](http://localhost:8080/SMA/Paginas/Information_retrieval/music-ir_org_virtual_home_of_music_information_retrieval_research.htm) 27 TFIDF: music mir workshop jcdl

[d:\Tomcat\webapps\SMA\Paginas\Information_retrieval/ir_detail.htm](http://localhost:8080/SMA/Paginas/Information_retrieval/ir_detail.htm) 22 TFIDF: inverter component compression lexical

Agente Login

- Responsável pela autenticação e criação de novos utilizadores;
 - Possui uma lista de agentes Interface disponíveis;
 - Cria agentes Interface quando não existe nenhum disponível.
-

Agente Interface

- Recebe pedidos do utilizador;
 - Devolve ao utilizador os resultados obtidos pelo Sistema Multi-Agente;
 - Analisa o comportamento do utilizador de modo a perceber os seus interesses;
 - Cria o perfil do utilizador;
 - Quando o utilizador entra numa página, dá sugestões de outras páginas interessantes.
-

Agentes de Análise

- Os diferentes agentes análise do SMA colaboram entre si na análise da lista de páginas recebidas do agente Interface;
- Aplicam o algoritmo TF-IDF;
- Existe um agente, Análise Coordenador, que coordena o trabalho de análise que está a ser efectuado;

Algoritmo TF-IDF

- Representa cada documento como um vector num espaço de vectores;
- Documentos com informação semelhante possuem vectores semelhantes;
- A estatística TF-IDF baseia-se no calculo:
 - Frequência de Termos $TF(w,d)$
 - Frequência nos Documentos $DF(w)$
 - Frequência inversa nos documentos $IDF(w) = \log(|D|/DF(w))$
- $TF-IDF(i) = TF(w_i,d) * IDF(w_i)$

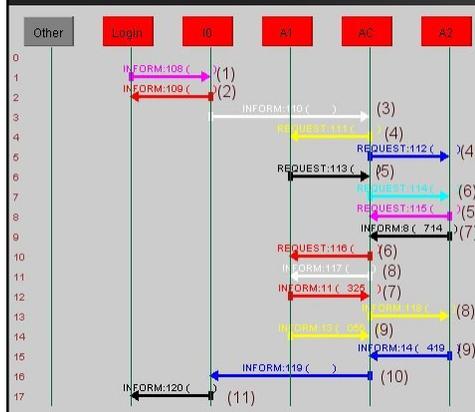
Agente Análise Coordenador

- Recebe do agente Interface pedidos de páginas a analisar e o respectivo critério de pesquisa:
 - Adiciona palavras-chave de acordo com o perfil do utilizador;
 - Distribui as páginas pelos agentes análise;
 - Recebe pedidos intermédios necessários para o calculo da estatística TF-IDF;
 - Ordena as páginas recebidas de acordo com o valor TF das palavras-chave.
-

Agente Análise

- Efectua as seguintes operações:
 - Remoção de código HTML;
 - Remoção de acentos;
 - Remoção de Palavras irrelevantes;
 - Extração de radicais;
 - Aplica o algoritmo TF-IDF;
 - Cria uma lista de palavras com o TF e TFIDF de cada palavra do documento;
 - Ordena o vector de palavras pela estatística TF-IDF
-

Troca de Mensagens



- 1-Utilizador Actual
- 2-Pedido Aceite
- 3-Lista de Páginas
- 4-Sub-Lista de Páginas
- 5-Pedido Cálculo DF (A – AC)
- 6-Pedido Calculo DF (Ac – A)
- 7-Calculo de DF Efectuado
- 8-Lista de Cálculos de DF
- 9-Solicitação Páginas Analisada (A–AC)
- 10-Solicitação Páginas Analisada (AC–I)
- 11-Saída de Sessão

Trabalho a desenvolver

- Efectuar análise semântica de modo a verificar termos de interesse nas páginas consideradas relevantes;