

Text Mining: a *Knowledge Based* em formato de texto

A Manchete (www.manchete.pt), a maior empresa Portuguesa de gestão e análise de informação pública, parte do grupo Novabase, apresentou este mês em Barcelona, no decorrer da 1ª Conferência de Utilizadores da Temis - Temis User Conference, o primeiro módulo de análise e investigação para a língua portuguesa: O TM360 Skill Cartridge (*Text Mining*).

O *Text Mining* surge de uma parceria para o desenvolvimento desta tecnológica excepcional, nos moldes em que está a ser implementada em Portugal. *“Ambas as empresas trabalharam em conjunto para poder desenvolver a análise de Texto analítico no contexto do mercado nacional, e para que existam bases no uso do texto em Português no mundo inteiro, foram as palavras de Charles Huot CEO e co-fundador da TEMIS, na User Conference.*

A Manchete, enquanto *platinum partner* da Temis, tem vindo a desenvolver esta plataforma de interpretação de dados disponível a partir de agora, através da área de Intelligence da empresa, e que permite estruturar toda a informação textual existente dentro das empresas clientes bem como a informação externa necessária para o desenvolvimento do negócio.

Esta mais recente aposta da Manchete desenvolve-se na área do conhecimento em prol das empresas e consiste num conjunto de técnicas e metodologias através das quais se pode focar e cruzar a informação recolhida (nas organizações) de forma a criar uma “rede” gráfica, baseada na identificação das entidades extraídas do texto. *“Vamos estruturar informação que nunca foi estruturada. Não ensinamos o computador a ler mas fazemos com que compreenda o que lê, identificando entidades como são o caso das pessoas, empresas, funções, produtos, valores, etc. e identifique as relações que estas estabelecem entre si.” afirma Pedro Rolim, Director de Inovação da Manchete e principal responsável pelo desenvolvimento do Text Mining em Portugal.*

Como resposta à mudança num mercado cada vez mais competitivo e global, as empresas e organizações sentem a necessidade de se prepararem para o embate que enfrentam. Essa preparação implica o conhecimento do meio onde se movimentam, focando particular atenção nas necessidades dos seus clientes e nas capacidades dos seus concorrentes.

A tendência do *Text Mining* será para crescer em Portugal como tem ocorrido noutros países, França é um bom exemplo, onde os sistemas de Business Intelligence já lidam com informação estruturada e não estruturada simultaneamente.

Com efeito, a quantidade de dados não estruturados disponíveis nunca foi tão grande como nos dias de hoje, tanto no âmbito interno como externo das empresas. *“Foi a pensar neste novo paradigma da gestão da informação e do conhecimento, em que os processos de tomada de decisão não podem permanecer fundamentados em informação numérica armazenada nos repositórios corporativos, que a Manchete reage e desenvolve em parceria esta nova fonte de conhecimento baseada no texto.”*, declara Fátima Rebelo, Managing Partner.

Na perspectiva de satisfazer a necessidade de informação relativamente ao ambiente empresarial envolvente, surgem cada vez mais empresas especializadas na recolha, tratamento e divulgação de informação num determinado sector, produzindo grandes volumes de informação. Segundo um estudo recente da Merrill Lynch e da Gartner referido por Delen e Crossland, 85% a 90% da informação corporativa é armazenada de forma não estruturada, ou seja, em texto, imagem, vídeo e áudio.

O probabilidade de transformar dados em informação e de desenvolver conhecimento sobre a actividade da empresa, o seu funcionamento ou os seus clientes a partir de dados armazenados pelas organizações ao longo de décadas ou a partir de informação gratuita, disponível e de âmbito interno e externo é extremamente apelativa e é a mais recente aposta da Manchete.

Manchete, a Knowledge based company.

Sobre o Text Mining

O *Text Mining* é considerado uma especialização dos processos de data mining confinada à informação textual, na medida em que esta metodologia visa igualmente a descoberta de conhecimento “escondido” nos repositórios de dados. Um dos principais, senão mesmo o principal factor de diferenciação entre data mining e text mining, é a tipologia da informação utilizada para alimentar os respectivos processos de análise. No caso do text mining a fonte de dados é o texto na forma de expressão natural por contraponto às bases de dados estruturadas usadas em data mining. O interesse crescente por processos de data mining e, mais recentemente, de text mining também se justifica pela possibilidade de aproveitamento da informação existente nos repositórios corporativos. Efectivamente, através destes processos torna-se viável utilizar informação que foi produzida pelos diversos sistemas das empresas com intuítos meramente operacionais e sobre a qual não era expectável a priori, ser usada para outros fins, para extrair conhecimento e, por conseguinte, gerar valor para a organização. Para além de dar nova vida à informação corporativa interna, os processos de text mining abrem as portas à utilização de novas fontes de conhecimento. Com a adopção destes processos, torna-se viável utilizar fontes de informação externa à organização, a maior parte dela gratuita ou com baixos custos de obtenção, para alimentação de processos de inteligência competitiva e marketing intelligence.

Sobre a Manchete

A Manchete é a maior empresa portuguesa na gestão e análise de conhecimento e informação pública do grupo Novabase. Conta já com inúmeras parcerias nacionais e internacionais e com a presença em Angola desde 2007. Com 13 anos de história e mais de 65 colaboradores, apresenta um volume de negócio na ordem do 1,9M€. Desenvolve a sua actuação segundo três orientações: *Media Based, Reputation Based e Intelligence Based*, a Manchete soma actualmente mais de 250 clientes nas mais variadas áreas de actividade, são disso exemplo a SAP, Deloitte, IBM, Pfizer, Somague, Ministério da Administração Interna ou Presidência do Conselho de Ministros. Desenvolveu o primeiro sistema de

Para mais informações, por favor contacte Filipa Pais de Sousa, +351 93 472 0061

M Public Relations, S.A. Espaço Castilho - Rua Castilho, 13 D- 7ºB 1250-066 Lx telf. +351 21 154 61 20



"Information Management" - MyNetpress - a funcionar em Portugal. A Manchete tem como principais factores de sucesso a aposta na qualidade dos serviços prestados, o acompanhamento personalizado dos clientes e um contínuo investimento na área da inovação, nomeadamente na Internet.

Para mais informações consulte o site www.manchete.pt

Sobre a TEMIS

A TEMIS é líder no desenvolvimento de soluções de software para a interpretação de Texto Analítico em empresas. O seu posicionamento visionário fica reconhecido com a solução: Luxid®, um software especialmente programado para responder às necessidades tanto na ciência, mundo empresarial, na edição, como até na segurança Interna. Conhecida também pelo desenvolvimento de uma forte capacidade táctica no uso da chamada "Intelligent Information" nas decisões estratégica da actividade empresarial, software como Competitive Intelligence, Scientific Discovery, Opinion Mining, Voice of the Customer and Content Publishing.

O Software Luxid® transforma dados não estruturados em conhecimento actual, possibilitando na análise de novos conteúdos uma descoberta de informação estratégica.

Fundada em 2000, a TEMIS opera nos Estados Unidos, França e Alemanha, e é representada no mundo inteiro através da sua rede de parceiros certificados.

TEMIS através das soluções/software inovadores têm sido a escolha pelas principais empresas/organizações, tais como: Agence France-Presse, BASF, Bayer Schering Pharma, BNP Paribas, Boehringer Ingelheim, CARMA Internacional, Convera, Editions Lefebvre-Sarrut, Elsevier, EMC, Europol, Ministério Francês da Defesa e das Finanças, Ingenuity, Investi n France Agency, Liquid Campaign, Merck Serono, Nature Publishing Group, Novartis, PSA Peugeot-Citroen, Roche Diagnostics, Sanofi-Aventis, Solvay Pharmaceuticals, Springer Science + Business Media, The McGraw-Hill Companies, Thomson Reuters. Para mais informações consulte www.temis.com

Para mais informações, por favor contacte Filipa Pais de Sousa, +351 93 472 0061

M Public Relations, S.A. Espaço Castilho - Rua Castilho, 13 D- 7ºB 1250-066 Lx telf. +351 21 154 61 20