

## IDÉIA DE PROJETO PROGRAMA IBEROEKA

Projetos em colaboração com empresas dos seguintes países: Argentina, Bolívia, Chile, Colômbia, Costa Rica, Cuba, El Salvador, Equador, Espanha, Guatemala, Honduras, México, Nicarágua, Panamá, Paraguai, Peru, Portugal, República Dominicana, Uruguai e Venezuela.

### DADOS DA EMPRESA

Nome da Empresa: upLexis Tecnologia S/S Ltda., CNPJ: 06.242.066/0001-74.

Endereço: Rua Luís Coelho, 340, Cj. 102, São Paulo – SP, CEP 01309-000, Brasil

Pessoa de Contato: João Marcelo Arcoverde

Telefone: (+55 11) 3255.3160

E-mail: jmarcelo@uplexis.com

Atividade da Empresa:

A upLexis desenvolve tecnologias especializadas na aquisição, organização, recuperação, tratamento e mineração de dados e textos para auxiliar processos decisórios em negócios corporativos;

Número de empregados:

- a) Faturamento no último ano fiscal:  
R\$ 3.938.263,98 (~ quatro milhões de reais)
- b) Faz parte de Grupo empresarial: Não

### RESUMO DA IDÉIA DE PROJETO E OBJETIVO

Mineração de dados na Web, ou simplesmente mineração Web, é o processo de descobrir informação útil em dados da Web, por meio de técnicas de mineração de dados. O custo de encontrar a informação correta não está associado à falta de técnicas de mineração de dados, mas à complexidade de gerenciar os dados Web necessários, e ao uso eficiente e eficaz de técnicas conhecidas de mineração.

Mineração Web é um processo iterativo, no qual prototipagem tem um papel essencial para experimentar facilmente com diferentes alternativas, bem como para incorporar o conhecimento adquirido durante iterações anteriores do processo. Para facilitar prototipagem, um nível apropriado de abstração precisa ser provido ao usuário ou programador responsável pela tarefa de mineração. O projeto aqui resumido representa uma das primeiras iniciativas para a criação de uma máquina de mineração para a Web.

Motivado pela falta de um arcabouço para abstração do processo de mineração Web conforme mencionado acima, foi desenvolvido em uma tese de doutorado um modelo para prototipagem intensiva em mineração Web, chamado de modelo WIM - Web Information Mining. O WIM já inclui um modelo de dados formal e uma álgebra, que estão especialmente projetados para manipular e minerar os principais tipos de dados presentes na Web. O objetivo final do WIM é reduzir significativamente o custo em minerar dados Web, de forma a fazer o processo de mineração mais acessível e de fácil uso.

WIM está atualmente implementado como um protótipo com uma parte das funcionalidades originalmente propostas. O protótipo ainda precisa ser transformado em uma ferramenta que possa ser usada mineração de dados tradicionais e dados Web, em escala industrial. Por uma visão mais prática, essa proposta inclui a extensão da interface do WIM, por exemplo para permitir que usuários possam graficamente usar o WIM como ferramenta no futuro. Por uma visão mais teórica, essa proposta inclui a extensão dos operadores do modelo, de forma que a cada dia mais tarefas de mineração de dados possam ser empregadas pelo WIM.

Ao final de seu desenvolvimento, espera-se que o WIM seja uma ferramenta facilitadora de desenvolvimento de aplicações em mineração de dados e mineração Web. Deve ser de fácil uso, para que usuários não programadores possam utilizar, e ainda deve suportar um conjunto significativo das possíveis tarefas de mineração de dados. Para o caso de mineração de dados Web, o WIM também propõe a coleta, indexação atualizada e organização de domínios da Web, como parte do processo de acesso aos dados em que propõe manipular.

#### PERFIL DE SÓCIO - EMPRESA BUSCADA

Existem dois perfis que interessam à **upLexis** para tornar o projeto do WIM viável. O primeiro perfil propõe o desenvolvimento do WIM em conjunto entre a **upLexis** e a empresa parceira em questão. Deve ser uma empresa de tecnologia da informação, preferencialmente que tenha seu negócio relacionado à área de Web, recuperação de informação, ou mineração de dados. Boa penetração no mercado também seria uma qualidade importante da parceira, para que a parte comercial também seja apoiada.

O segundo perfil propõe o desenvolvimento totalmente na **upLexis**. A empresa parceira deve aportar capital, e preferencialmente ter a capacidade de contribuir para a inserção no mercado do novo negócio.