

alpha2.txt

ALPHA/2: Fonetizador para a Língua Portuguesa

Manual de Apresentação e Utilização

Recuperação de Informações

A utilização de identificadores numéricos, aos quais se pode acrescentar dígitos verificadores. Em diversos serviços relacionados a busca de informações, como Auxílio a Lista nas Companhias. O uso de arquivos organizados em ordem alfabética é bastante comum em repartições públicas, como Na sociedade informatizada, porém, as informações estão armazenadas em meios magnéticos e sua Isso torna bastante complexos os processos de acesso a dados a partir de nomes, que podem ter

Codificação Fonética

Uma forma de tornar viável a recuperação de informações a partir de nomes é pelo método da codificação fonética. Como a função fonética não poderia ser biunívoca, desde quando nomes diferentes podem levar a Na verdade, o método de fonetização não tem a pretensão de incorporar a capacidade humana de a A Torre de Babel

Muitas pesquisas e experimentos foram realizadas na busca de um sistema de fonetização de nomes. A partir daí, no decorrer dos séculos, os homens afastaram-se mais ainda, criaram nações e com Em função disso, o fonetizador depende da língua do país onde é usado, mesmo que gere códigos ALPHA/2: Fonetizador da Língua Portuguesa

A ALPHA/2 é uma nova versão aperfeiçoada e portátil da ALFONET, um fonetizador que foi utilizado Nas experiências efetuadas com operadores brasileiros, a fonetização obtida foi perfeita não s

- o W inglês soa como U;
- o J espanhol e o H inglês soam como R;
- CC em italiano lembra o nosso CH;
- etc.

Tal fato não chega a atrapalhar, pois é normal que os atendedores solicitem a pessoas com nomes A ALPHA/2 incorpora uma função especial, que é o tratamento opcional de sinônimos, o que torna Caracteres válidos para fonetização

- Letras maiúsculas não acentuadas (A - Z);
- Letras minúsculas não acentuadas (a - z);
- Algarismos arábicos (0 - 9), desde que formem um número isolado (ex.: 123) ou precedam
- Espaços em branco;
- Caracteres especiais, que são eliminados durante o processo:
 - Ponto (.);
 - Vírgula (,);
 - Hífen (-);
 - Barra (/);
 - E comercial (&);
 - Apóstrofe (')

É possível ainda, através da elaboração de uma "rotina filtro", serem admitidos os caracteres

Tal "rotina filtro" transformaria as vogais acentuadas nas correspondentes vogais não acentuadas

Exemplos de Fonetização

Os seguintes grupos de nomes possuem os mesmos códigos fonéticos:

- Luiz Eduardo Siqueira e Luis Eduardo Cerqueira
- Virgildasio Sena, Vigidazio Cena, Viljidarzio Senna
- Jorge Figueira e George Filgueiras
- MarcoAntonio e Mark Antony
- Joji Nakuma e Jorge Nacuma
- Farmácia Santana e Drogaria Sant´ana
- Churrascaria Xanadu e Restaurante Chanandum
- Associação Educacional Baiana e Grêmio Cultural da Bahia
- Afonso Imóveis e Alfonso Imobiliária
- etc.

Utilização

A ALPHA/2 pode ser chamada a partir de programas escritos em qualquer linguagem que possua su

Em qualquer dos ambientes, a chamada a ALPHA/2 é feita através da passagem de 3 parâmetros:

1. Tamanho do nome a fonetizar:

- ASSEMBLER: TAMNOME DS H

- COBOL: Tamanho-nome PIC 9(3) COMP.
- NATURAL: TAMANHO-NOME B(2)

2. Código de comunicação:

O programa usuário envia 0 para fonética pura ou 1 para fonética com sinônimos.

A ALPHA/2 devolve 0 se a codificação fonética foi normal, 1 no caso de ocorrência de caracter

- ASSEMBLER: COMUNIC DS CL1
- COBOL: Codigo-comunicacao PIC X.
- NATURAL: CODIGO-COMUNICACAO A(1)

3. Nome a fonetizar e código fonético

O programa usuário envia o nome composto a fonetizar, recebendo da ALPHA/2 o código fonético

O código fonético gerado será constituído por segmentos de 4 bytes, cada um deles corresponden

- ASSEMBLER: NOMEFON DS CLnnn
- COBOL: Nome-e-fonetico PIC X(nnn).
- NATURAL: NOME-E-FONETICO A(nnn)

Chamada a ALPHA/2:

- ASSEMBLER: CALL ALPHA2, (TAMNOME, COMUNIC, NOMEFON)
- COBOL: CALL 'ALPHA2' USING Tamanho-nome
Codigo-comunicacao
Nome-e-fonetico
- NATURAL: CALL 'ALPHA2' TAMANHO-NOME
CODIGO-COMUNICACAO
NOME-E-FONETICO

Exemplo:

O programa usuário envia a ALPHA/2 os seguintes parâmetros:

- Tamanho-nome = 48
- Codigo-comunicacao = 0
- Nome-e-fonetico = LUIZ EDUARDO DE OLIVEIRA SIQUEIRA

A ALPHA/2 retornará

- Tamanho-nome = 48
- Codigo-comunicacao = 0
- Nome-e-fonetico = L2__EDD3OLVRSCR2_____

Pesquisa Fonética

Recuperar as informações armazenadas a partir de chaves fonéticas depende da elaboração de rotinas.

O que se sugere é a implementação nessas rotinas de vários níveis de pesquisa:

- 1o. nível: chave fonética completa.
- 2o. nível: chave fonética genérica com o primeiro segmento e obrigatoriedade de h
- 3o. nível: chave fonética genérica com o primeiro segmento e obrigatoriedade de h
- 4o. nível: chave fonética genérica com o primeiro segmento e existência de pelo m
- etc.

Exemplo:

- PEDRO VIEIRA DA SILVA FERNANDEZ está no cadastro a ser pesquisado.
- O nome disponível para a pesquisa é PEDRO DA SILVA VIEIRA.

O resultado positivo da pesquisa só será atingido no 3o. nível.