



[INTERNET]

TUMBA! TOMBA

No mundo virtual, há que guardar hoje para ler amanhã

ALEXANDRA ROSA

Google, Yahoo, AltaVista, Accoona, Ask.com, MSN, AOL, Netindex. Quem procura informação na Internet de certeza que já usou um destes motores de busca. Mas há pessoas que não se satisfazem em ser simples utilizadores de ferramentas informáticas e querem perceber como é que estas funcionam. É o caso de alguns investigadores do Laboratório de Sistemas de Informação de Grande

Escala (LaSIGE), sediado na Faculdade de Ciências de Lisboa (FCUL), que, em 2001, começaram a desenvolver um buscador para a Web portuguesa. Ao resultado chamaram Tumba! – o acrónimo de *Temos um Motor de Busca Alternativo!* «Queríamos aprender coisas novas», afirma Daniel Gomes, 29 anos, um dos engenheiros informáticos que, desde o início, integrou o projecto. Na verdade, o Tumba!, embora coordenado pelo investigador sénior Mário Silva, foi implementado por estudantes que, paralelamente, trabalhavam

nos seus projectos de mestrado ou doutoramento. «Saiu-nos do corpo», recorda Daniel. Afinal, uma das «coisas» que a equipa aprendeu foi que criar e manter um motor de busca não é nada fácil.

PARA CONSTRUÍREM o Tumba!, os carolas informáticos da FCUL começaram por identificar e guardar páginas portuguesas de Internet. Destes documentos, extraíram todas as palavras e formaram um índice invertido – o equivalente *internético* ao índice remissivo dos livros –, a partir do qual

são localizadas rapidamente as páginas Web que contêm as palavras-chave procuradas pelo utilizador.

Depois, foi «só» desenvolver um algoritmo computacional para ordenar, por importância, os documen-

pela quantidade de hiperligações (*links*) recebidas. No entanto, a construção de um bom algoritmo de ordenação mistura geralmente várias estratégias secretas, ou não fosse este um dos pontos-chave para o sucesso do negócio.

AO FIM DE QUATRO ANOS, o *www.tumba.pt* estava pronto. Mas hoje, para funcionar, precisaria de manutenção e esta tornou-se cada vez mais difícil desde que os «pais» do projecto foram completando os seus graus académicos e abandonando o laboratório da FCUL.

Porém, conforme foram armazenando páginas, «até os computadores rebentarem pelas costuras», Daniel e os seus colegas aperceberam-se do enorme património sociológico que guardavam. Actualmente, as ferramentas Web permitem a produção de um número imenso de documentos virtuais – retratos quotidianos dos nossos tempos –, os quais não chegarão (como os livros) às gerações futuras porque apenas estão disponíveis efemeramente. «É História que se perde todos os dias», alerta Daniel. Assim, nasceu a ideia de criar uma espécie

Os portugueses criadores do motor de busca Tumba! SOMARAM-LHE uma Torre do Tombo virtual – a Tomba

tos encontrados. É que qualquer cibernauta quererá obter a informação que procura o mais depressa possível, sem ter de ler muitos textos. Por isso, num motor de busca eficaz, os sítios *www* mais relevantes para a pesquisa deverão aparecer «à cabeça». Uma das formas típicas para calcular a importância de uma página é a indexação pelo número de vezes que é citada. Ou seja,

de Torre do Tombo virtual – a Tomba, um subproduto do Tumba!

Para já, o primeiro repositório de Internet portuguesa, realizado com o apoio da Fundação para a Computação Científica Nacional, conta com mais de 57 milhões de documentos arquivados. Assim, no futuro, com a Tomba, ninguém poderá dizer: «Porque é que não guardei isto!» ■