

WEBPAPER – MÁS PERGUNTAS E BOAS RESPOSTAS: A ARTE DE INTERROGAR

Alberto Simões, Paulo Rocha, Rúben Fonseca

Resumo: Neste documento apresentamos o conceito de WebPaper, a resposta a um conjunto de questões de vários graus de dificuldade usando como recurso base a Internet. Em primeiro lugar é discutido o conceito de WebPaper e o tipo de questões que são colocadas. Em seguida são descritas as experiências já realizadas e em que contextos. Finalmente, apresentamos uma *framework* para a ajuda na realização de WebPapers.

Abstract: In this document we present the concept of WebPaper, a question-answering contest for humans using the Internet as a resource. First, we discuss the concept of WebPaper and the kind of questions used. Second, we describe the experiments performed and their contexts. Finally, we present a framework for the creation of WebPapers.

1. WebPaper: Conceito

O conceito do WebPaper surge originalmente na obra de ficção científica *As Fontes do Paraíso* (Clarke 1979), onde o autor descreve uma competição algo semelhante, que denomina de *retrieval competition*. Com esta ideia em mente e com a necessidade de organizar concursos pedagógicos inseridos em eventos organizados pelo Departamento de Informática da Universidade do Minho, levou à criação de uma competição a que chamamos WebPaper. Assim como num Rally Paper os concorrentes andam de carro a tentar responder a perguntas, e num Pedipaper os concorrentes andam a pé, num WebPaper os concorrentes andam na Web, à procura de respostas.

Embora a medição à priori do grau de dificuldade das perguntas seja subjectivo, optámos por, no geral, dar diferentes pesos às perguntas apresentadas, de acordo com os graus de indirectão que são necessários percorrer desde a pergunta até se chegar à resposta correcta.

O tipo mais simples de pergunta, e bastante frequente nos primeiros WebPapers, é aquela que apresenta uma resposta imediata, e que pode ser comparada às perguntas típicas do *Trivial Pursuit*, como sejam, Qual a capital do Arkansas? ou Qual a principal personagem de «O Idiota»? Este tipo de perguntas, no entanto, torna-se cada vez menos interessante devido à expansão da Wikipedia (Wikipedia 2007), onde as respostas podem ser facilmente encontradas.

Modos de obrigar os participantes a procurar para além da Wikipedia, usados nos mais recentes Webpapers tem sido o uso de perguntas com maior número de indirectões (Qual a capital do quarta província mais populosa do Canadá?), e perguntas enredadas (Quem ganhou a Taça da Suíça no ano de nascimento do ministro das finanças britânico?).

Um tipo de perguntas usado desde o início, mas que tem vindo a ser cada vez mais usado, são as perguntas que não contêm nomes próprios raros (que levariam a uma resposta imediata mediante a sua pesquisa na Web). Repare-se na diferença entre estas duas perguntas: Quem foi a única candidata a um Óscar representando um papel masculino? e Em que filme foi Linda Hunt candidata a um Óscar por um papel masculino?

A segunda pergunta exige apenas uma busca pelo nome da actriz – provavelmente na Wikipedia ou no *International Movie Database* (Internet Movie Database 2007) – que conduzirá certamente a uma biografia indicando este factóide. Por sua vez, a primeira pergunta exige a criação de expressões de pesquisa mais elaboradas, pelo menos se o participante quiser evitar percorrer uma lista de cerca de 700 candidaturas aos prémios de Melhor Actriz e Melhor Actriz Secundária.

Outro tipo de perguntas em expansão são as perguntas que, devido às suas características, não podem ser encontradas na Wikipedia: Qual o preço de um bilhete de comboio de Salamanca a Bilbao? ou Qual o endereço da embaixada portuguesa na África do Sul?

Por vezes, outras perguntas são feitas de forma a que o concorrente seja induzido em erro, como por exemplo com: Com que se casou a rainha interpretada no cinema por Greta Garbo? Se foi banal descobrir que a pergunta se referia à rainha Cristina da Suécia, os participantes perderam imenso tempo tentando descobrir com quem a rainha se casou, uma vez que poucas páginas referiam explicitamente o seu celibato; estas perguntas têm uma taxa de resposta muito baixa, o que leva a pensar que os participantes têm dificuldade em concluir que estão a partir de uma premissa falsa.

Outro tipo interessante de perguntas é as perguntas baseadas numa imagem: por exemplo, identificar uma bandeira pouco conhecida, o bairro de Oslo onde determinada estátua poderia ser encontrada, ou o significado de determinado carácter chinês.

Algo já sugerido, mas nunca testado, é o uso de perguntas com som (por exemplo, identificar uma música), devido naturalmente à impossibilidade de usar um motor de pesquisa para descobrir uma música

trauteada.

Quanto à língua em que as perguntas são efectuadas, sempre se assumiu um domínio da língua inglesa pelos participantes uma vez que a Web em português e brasileiro são e sempre foram mais limitadas no seu âmbito. Sendo assim, nunca se omitiu uma pergunta por esta não ter resposta em português na Web.

No que diz respeito à avaliação, sempre aconselhámos a apresentação de um URL contendo a resposta, para confirmar respostas distintas ou para prever o caso de um erro da organização.

2. Experiências Realizadas

O primeiro WebPaper realizado foi em 2003, nas Jornadas de Informática da Universidade do Minho (JOIN'03). Desde esse ano até 2006 que todas as Jornadas de Informática têm vindo a acolher a organização de um WebPaper. Cada ano o WebPaper tem vindo a aumentar em número de participantes incluindo não só alunos das licenciaturas que organizam o evento (Licenciatura em Matemática e Ciências da Computação e Licenciatura em Engenharia de Sistemas e Informática) mas também convidados (palestrantes) das Jornadas. No ano de 2006 obtivemos um recorde de participações, com 30 grupos.

Depois da sessão habitual de duas horas em que os grupos tentam encontrar respostas às perguntas usando todos os meios disponíveis na Internet, foram sempre realizadas sessões de encerramento onde as perguntas são apresentadas e as respectivas respostas. Algumas das respostas erradas são escolhidas para serem divulgadas na sessão de encerramento (embora sem se divulgar o grupo que as respondeu) como forma de entretenimento, mas também para corrigir algumas falsas ideias dos participantes.

Além destes quatro WebPapers realizados com um conjunto equilibrado de participantes, em 2004 o conceito foi usado durante a Minho Campus Party 2004, realizada em Braga, e com a participação de alunos de várias faixas etárias. Devido ao grande número de participantes (cerca de 60 grupos) e devido à falta de um sistema automático de correcção (o que apresentamos na secção seguinte é mais recente), foi utilizado um sistema semi-automático para uma correcção mais eficiente, mas que não permitiu a análise das respostas erradas.

O Departamento de Informática recebe visitas de escolas secundárias durante todo o ano lectivo para apresentação das suas áreas de investigação. Nessa visita é habitual inserir-se um espaço lúdico onde até há bem pouco tempo os alunos se divertiam com jogos first person shoot que pouco ou nada têm de educativo. Ultimamente foi adaptado o conceito do WebPaper, tendo em atenção a faixa etária e o pouco tempo (cerca de 20 minutos) disponível. Neste sentido, criou-se uma base de dados de perguntas, e um conjunto de 4 respostas, das quais apenas uma está correcta, de modo a que por um lado seja mais simples a resolução do WebPaper, e por outro, a correcção possa ser feita de forma totalmente automática.

Alguns docentes do Departamento de Informática adaptaram alguns dos webpapers utilizados nas Jornadas de Informática para serem propostos a alunos de outros cursos, como forma de incentivo ao estudo e importância de técnicas de pesquisa na Internet.

3. Interface Web

O interface Web utiliza uma série de tecnologias para facilitar a realização de WebPapers, tanto do lado do cliente (concorrente), como do lado do administrador (organizador). Em primeiro lugar, todo o interface foi desenhado usando standards Web (XHTML e CSS) e uma framework de criação de aplicações Web de nome Catalyst. Esta framework é escrita na linguagem de programação Perl, o que permite uma extrema flexibilidade e robustez.

Para armazenar o estado da aplicação usamos uma base de dados relacional usando o esquema indicado na figura seguinte. Esta base de dados permite-nos guardar toda a informação sobre a aplicação de uma forma extremamente eficiente e evitando redundância de dados.

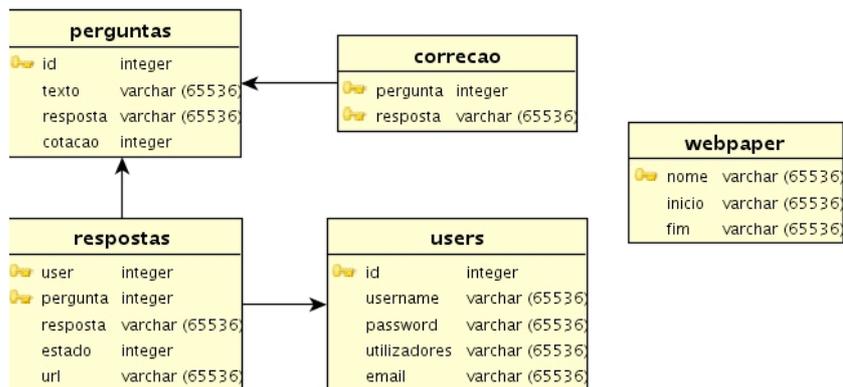


Figura 1: Esquema relacional da base de dados.

A interface está separada em duas secções distintas: uma que permite efectuar toda a administração do sistema, e outra que permite a utilização do mesmo para responder às questões formuladas. Iremos analisar pormenorizadamente cada uma destas secções.

3.1 Administração

A secção de administração, a que chamaremos de backend, é acedida e protegida através de uma password de segurança (que poderá ser comodamente alterada a qualquer altura). Depois de autenticado, o administrador poderá ver o seguinte ecrã, que apresenta informações úteis em relação ao progresso do seu WebPaper.

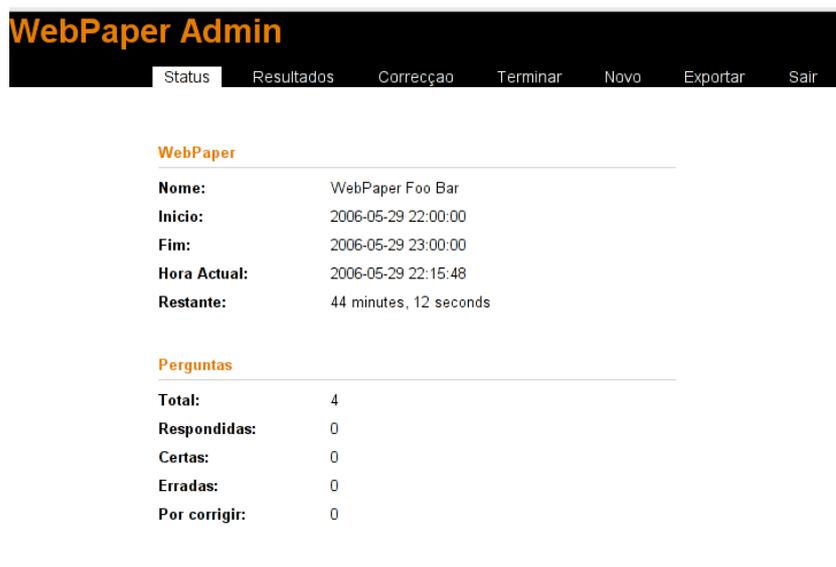


Figura 2: backend de administração do WebPaper

Como é visível na figura, além de informações úteis sobre o estado do WebPaper a decorrer, a área de backend permite uma série de operações de controlo sobre o mesmo, como sejam:

- Status Permite ver o estado actual do WebPaper
- Resultados Mostra o resultado por ordem decrescente de pontuação de todos os grupos, depois de todas as perguntas serem corrigidas
- Correção Entra no modo do correção assistida de perguntas.
- Terminar Termina prematuramente o WebPaper
- Novo Inicia um novo WebPaper
- Exportar Exporta o WebPaper actual (perguntas e dados de inicialização) para um formato portátil (XML) que pode ser novamente importado no menu Novo
- Exportar Sai do modo de administração

3.1.1 Criação de um WebPaper

Para facilitar a criação de todo o webpaper, incluindo datas e horas de início e término, perguntas e cotações, a aplicação está preparada para receber um ficheiro no formato XML com todos os dados necessários para a realização do mesmo. O XML é um formato de texto convencional, com alguma sintaxe definida de modo a descrever de uma maneira clara os dados nele contidos.

De seguida, segue-se um exemplo simples do que seria um WebPaper definido neste formato:

```
<?xml version='1.0' standalone='yes'?>
<webpaper>
  <inicio>2006-04-05 17:00:00</inicio>
  <fim>2006-04-05 18:30:00</fim>
  <nome>WebPaper JOIN'06</nome>
  <pergunta>
    <texto>Qual a quarta província mais populosa do Canadá?</texto>
    <resposta>Alberta</resposta>
    <cotacao>15</cotacao>
  </pergunta>
  <pergunta>
    <texto>Quem foi a última pessoa conhecida a falar etrusco?</texto>
    <resposta>o Imperador Cláudio</resposta>
    <cotacao>5</cotacao>
  </pergunta>
  ...
</webpaper>
```

Com este ficheiro criado, iríamos à opção "Novo" conforme a figura seguinte indica, e introduziríamos o nosso ficheiro que seria automaticamente processado, inicializando assim o WebPaper.



Figura 3: Criação de um WebPaper.

3.1.2 Correção de um WebPaper

Esta fase é a mais delicada de todo o processo. O objectivo seria que o computador conseguisse corrigir automaticamente todas as respostas dos concorrentes, indicando de forma acertada quais as correctas e quais as erradas. No entanto, isto não é possível por vários motivos.

Primeiro, o utilizador poderá não inserir exactamente o texto que é definido como resposta correcta. Poderá até mesmo colocar "lixo" antes ou depois da resposta. Por exemplo, definindo como resposta oficial: "Rua de S. Barnabé" qualquer resposta como "A rua chama-se Rua de São Barnabé" não iria ser aceite por não ser totalmente igual à resposta modelo.

Por outro lado, existem situações em que mais que uma resposta é possível, muitas vezes pela pergunta ter sido mal formulada, ou até mesmo por na Web existirem fontes contraditórias sobre a resposta à pergunta.

Assim, torna-se necessário usar um processo de correcção assistida e não automática. O processo funciona da seguinte maneira:

- O programa corrige automaticamente todas as respostas que sejam iguais às respostas definidas no ficheiro XML (a comparação é feita em minúsculas, e depois de limpar espaços em branco a mais no início e no fim da resposta).
- As respostas assim não corrigidas são revistas pelo administrador que, uma por uma, indica se a resposta dada pelo utilizador está ou não correcta.

- Caso a resposta seja marcada como correcta, todas as respostas de outros utilizadores que sejam iguais a esta nova resposta, são também marcadas como correctas.
- Caso a resposta seja marcada como errada, todas as respostas de outros utilizadores que sejam iguais a esta, são marcadas como erradas.
- O processo repete-se até que todas as perguntas tenham sido corrigidas.

Como se pode ver, o processo de correcção assistida poderá facilitar bastante a árdua tarefa de corrigir todas as perguntas, poupando tempo ao administrador e obtendo os resultados finais mais rapidamente. Segue-se uma figura deste processo em acção.



Figura 4: Processo de correcção do WebPaper

No final, depois de todas as perguntas corrigidas, os resultados poderão ser listados conforme a figura seguinte.



Figura 5: Listagem de classificações

Ao clicar no nome do grupo (neste caso, apenas um) serão fornecidas todas as informações de registo deste grupo, com vista a identificar o mesmo. Assim, através desta área de backend, todo o processo do WebPaper pode ser controlado e administrado facilmente, com apenas alguns cliques.

3.2 Participação

Toda a área de participação no WebPaper, a que chamamos de frontend, tem como objectivo fornecer o interface necessário para os concorrentes se registarem e responderem de uma maneira cómoda às perguntas fornecidas. Assim, e apenas depois da hora de início do WebPaper definido na área de administração, o sistema permite aos utilizadores entrarem ou efectuarem o registo do seu grupo, conforme a seguinte figura.

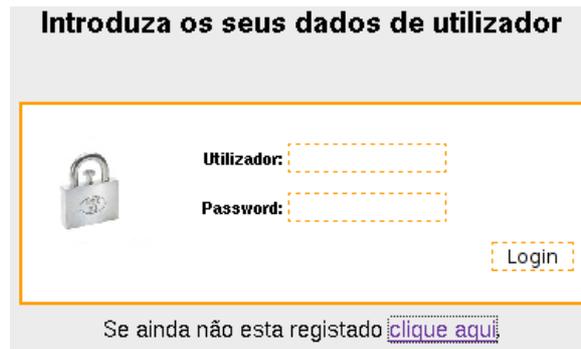


Figura 6: Página de login no WebPaper.

Depois de efectuar o registo o utilizador pode então autenticar-se com o sistema, permitindo assim o acesso às perguntas. A interface para a resposta das perguntas é extremamente intuitivo.

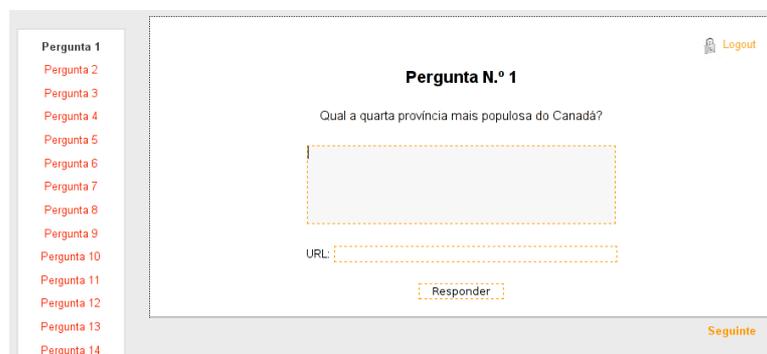


Figura 7: Interface do concorrente no WebPaper.

A lista das perguntas a responder é colocada do lado esquerdo. Do lado direito aparece a pergunta actual a ser respondida, com um campo para a resposta e outro para colocar (opcionalmente) o URL do site onde a resposta a esta pergunta foi encontrada.

Ao seleccionar o botão “Responder”, o sistema regista a resposta do utilizador na base de dados e avança automaticamente para a próxima pergunta não respondida, iterando o processo até que todas sejam respondidas. A qualquer momento (desde que esteja dentro do tempo regulamentar) o utilizador poderá alterar uma resposta a qualquer pergunta que já tenha respondido anteriormente.

As perguntas respondidas são marcadas a verde na barra lateral esquerda; as ainda não respondidas a vermelho, e a pergunta actual aparece seleccionada a preto. O utilizador poderá também facilmente sair do WebPaper através do botão “Logout”.

Quando o sistema detecta que acabou o tempo disponível deixa de aceitar mais respostas por parte dos utilizadores, terminando assim com a sua participação.

4. Conclusões

O WebPaper é um divertimento, quer para quem concorre quer para quem organiza. No entanto, e embora seja possível (como vimos) a criação de um sistema de ajuda à manutenção do WebPaper não é nada simples criar as perguntas. Estas além de não poderem ser de resposta imediata têm de, de alguma forma, entreter o concorrente.

Embora tenha sido realizado pela primeira vez como um passatempo para estudantes universitários, pode ser facilmente adaptado a aulas, e a outras finalidades educativas. No entanto, e mesmo na sua vertente lúdica continua a ser um excelente meio de aprender e fazer aprender.

A criação de um interface para a realização de WebPapers permitiu que mais tempo fosse investido que realmente torna o WebPaper interessante: as perguntas. Em relação ao sistema, além de fácil de usar é robusto. Embora de momento a criação de um WebPaper obrigue ao uso do formato XML, está em curso um melhoramento na interface de administração para se manter uma base de dados de perguntas que possam vir a ser reutilizadas de concurso a concurso, e mesmo para permitir a geração automática de WebPapers de forma aleatória.

Referências

Arthur C. Clarke (1979) *The Fountains of Paradise*.

Catalyst Web Framework (2006). <http://www.catalystframework.org/>, (consultado na Internet em Maio de 2006).

Internet Movie Database. (2007, March 15). In Wikipedia, The Free Encyclopedia. (consultado na Internet em Março de 2007 em

http://en.wikipedia.org/w/index.php?title=Internet_Movie_Database&oldid=115436614)

Perl.com, The Source for Perl (2006). <http://www.perl.org/>, (consultado na Internet em Maio de 2006).

Wikipedia. (2007, March 16). In Wikipedia, The Free Encyclopedia. (consultado na Internet em Março de 2007 em <http://en.wikipedia.org/w/index.php?title=Wikipedia&oldid=115539371>)

XML (2006). <http://en.wikipedia.org/wiki/XML> (consultado na Internet em Maio de 2006).