

UM ASSISTENTE DIGITAL PARA RESPONDER AUTOMATICAMENTE PERGUNTAS DE USUÁRIOS HUMANOS EM PORTAIS CORPORATIVOS

Lizandro Kirst da Silva
Daniel Ribeiro Brahm
Gustavo Tagliassuchi
Stanley Loh

ADS Digital Ltda.
www.adsdigital.com.br

lizandro{a}adsdigital.com.br
daniel{a}adsdigital.com.br
gustavo{a}adsdigital.com.br
stanley{a}adsdigital.com.br

Resumo

O presente trabalho procura minimizar os problemas relativos à busca de informações em Portais Corporativos com o uso de um assistente digital (AD). Um AD é uma ferramenta de auxílio ao usuário capaz de encontrar informações relacionadas a um assunto de maneira rápida, intuitiva e eficaz. Um AD interage com usuários humanos através de um navegador Web (*browser*), recebendo perguntas textuais e respondendo automaticamente (sem intervenção humana) a estas perguntas com base em conhecimento previamente armazenado. A principal vantagem do uso de AD's em Portais de Gestão do Conhecimento está no ganho de tempo, uma vez que o usuário encontra a informação num único local (não precisa passar por várias páginas Web) e utilizando sempre a mesma forma (perguntas em linguagem natural).

Este artigo apresentará a funcionalidade do AD proposto, bem como as técnicas empregadas para sua implementação. Também serão discutidas aplicações de uso e apresentado em detalhes um estudo de caso onde o assistente já está sendo utilizado.

Palavras-chave: assistente digital, interfaces inteligentes, portais corporativos, chatterbots, interação em linguagem natural

Introdução

O conhecimento de uma organização é formado a partir do conhecimento dos seus membros e pode aumentar à medida que este conhecimento é armazenado e compartilhado. Este é o conceito de organizações que aprendem (*learning organization*), conforme SENGE (2001). Para auxiliar na organização deste conhecimento, surgiu a área de Gestão do

Conhecimento (GC). A GC tem seu foco voltado para questões relacionadas a como as organizações podem tirar maior proveito do conhecimento existente dentro delas, facilitando a distribuição deste conhecimento entre seus membros, encorajando o registro das soluções adotadas para resolução de problemas, evitando a perda do conhecimento de seus especialistas quando estes deixam a organização e gerando novos conhecimentos (DAVENPORT & PRUZAC, 1997).

Uma das soluções tecnológicas mais utilizadas hoje em dia para suportar a Gestão do Conhecimento são os Portais Corporativos. Estes portais são sites Web especialmente projetados para o armazenamento e a recuperação do conhecimento disponível em uma organização. Uma das vantagens é permitir que pessoas divulguem o conhecimento que adquiriram e outras pessoas possam reutilizar este conhecimento. Com os portais, evita-se que o conhecimento se perca quando alguém sai da organização ou que a produtividade seja prejudicada quando colegas interrompem o trabalho de outros para pedir ajuda, justamente quando o conhecimento requerido encontra-se disponível de forma explícita.

Com o crescimento e popularização das tecnologias Web, ficou fácil usar portais corporativos, o que aumentou o volume de informações disponíveis. Apesar dos grandes benefícios, tal fato teve como consequência negativa a sobrecarga de informações (*information overload*), que acontece quando o usuário tem muita informação ao seu alcance, mas não tem condições de encontrar o que realmente deseja ou lhe interessa (CHEN, 1994).

Para evitar que usuários fiquem “perdidos” em meio a tantas informações, os portais corporativos podem conter uma lista de FAQ’s (*frequently asked questions*), ou seja, uma lista com as perguntas mais frequentes e as respostas para estas indagações. Tais perguntas ajudam a encontrar mais rapidamente áreas específicas do site, informações sobre produtos e serviços ou mesmo conhecimentos mais complexos (por exemplo, como realizar determinado processo). As FAQ’s evitam, por exemplo, que especialistas humanos sejam requeridos a toda hora para dirimir dúvidas (por exemplo, em um *Call Center*).

Entretanto, quando a lista de FAQ’s é muito grande, fica difícil para o usuário encontrar a pergunta mais adequada a seu problema. Para verificar um exemplo deste problema, basta acessar a lista de FAQ’s com informações relativas à declaração do imposto de renda no site da Receita Federal (www.receita.fazenda.gov.br).

O presente trabalho procura minimizar os problemas relativos à busca de informações em Portais Corporativos com o uso de um assistente digital (AD). Um assistente digital é uma ferramenta de auxílio ao usuário capaz de encontrar informações relacionadas a um assunto de maneira rápida, intuitiva e eficaz. A interação entre usuário e o AD deve acontecer de maneira natural como se fosse num *Web chat*.

Um AD deve ser capaz de interagir com usuários humanos de um Portal Corporativo (ou de qualquer site) através de um navegador, recebendo perguntas textuais e respondendo automaticamente (sem intervenção humana) a estas perguntas com base em conhecimento previamente armazenado.

Este artigo apresentará a funcionalidade do assistente digital proposto, bem como as técnicas empregadas para sua implementação. Também serão discutidas aplicações de uso e apresentado em detalhes um estudo de caso onde o assistente já está sendo utilizado.

Assistentes Digitais e a Gestão de Conhecimento

Sistemas que respondem perguntas de usuários humanos existem há tempos. Eles são chamados de “sistemas conversadores ou de conversação”. O mais famoso é o sistema Eliza (WEIZENBAUM, 1967), que simulava uma seção de psicanálise. Na Internet, podem ser encontrados diversos destes sistemas (na Internet, são chamados “*chatterbots*”).

O mais famoso aqui no Brasil, não só pelo nível de conhecimento, quanto pela interação conseguida com o usuário, é a Sete Zoom (www.setezoom.com.br). Sete Zoom é um *chatterbot* com uma personagem virtual feminina. Ele foi criado pela Insite para o site da Close Up (Gessy Lever). Sua função é interagir com os internautas que acessam o site, mantendo uma conversação em linguagem natural. Este *chatterbot* se tornou famoso por ser um modelo virtual feminino e por sua personalidade atrevida e despojada. Entretanto, este conversador não tem por objetivo passar informações sobre produtos, mas tão somente fazer *branding*, ou seja, tornar conhecida e fortalecer uma marca.

Um porém dos sistemas conversadores é que seu objetivo é apenas o de manter a conversação de forma coerente, fingindo ser um interlocutor humano ou inteligente. Apesar da tecnologia utilizada e dos bons resultados, estes sistemas não são adequados para fornecer informações.

Já os assistentes digitais procuram responder às questões de usuários com coerência, mas o foco está no fornecimento de informações. Eles são especialmente úteis para usuários leigos na Web ou que estejam perdidos em um Portal. Assim, é mais fácil encontrar a informação desejada perguntando a alguém do que procurando na estrutura de *hiperlinks* do site. Não é objetivo do assistente simular ou imitar um interlocutor humano mas tão somente fornecer informações. Dentro deste objetivo, um AD não precisa responder a todo tipo de pergunta (por exemplo, sobre cultura geral). A imitação humana pode até prejudicar a atividade de busca de informações, uma vez que interações humanas contêm interjeições, frases incompletas ou com conotações, variações bruscas de assuntos, etc. (ZUE, 1997).

Um exemplo de assistente digital é discutido em ZUE (1997). O sistema Júpiter permite a conversação por telefone sobre a previsão do tempo em algumas cidades. O sistema possui um vocabulário de 1500 palavras. Primeiro, o sistema reconhece palavras faladas, depois monta um contexto e gera uma paráfrase em SQL (linguagem padrão para consultas a bancos de dados relacionais). De posse de um comando SQL, o sistema consulta um banco de dados e gera a locução correspondente da resposta.

Portais de Gestão do Conhecimento podem fazer uso de AD's para recuperação de conhecimento. A principal vantagem está no ganho de tempo para procura, uma vez que o usuário encontra a informação num único local (não precisa passar por várias páginas Web) e utilizando sempre a mesma forma (perguntas em linguagem natural).

Da mesma forma, o armazenamento do conhecimento é facilitado, já que os AD's permitem o armazenamento de forma padrão, através de perguntas e respostas. Isto evita a diversidade de formatos para armazenamento de conhecimento hoje disponíveis, tais como bancos de dados, e-mails e listas de discussão, páginas Web, FAQ's, documentos eletrônicos, etc. Esta diversidade é ruim porque confunde quem armazena e quem procura informações.

Assistente digitais podem ser utilizados em diversas aplicações, dependendo apenas do conteúdo armazenado na base de conhecimento. A seguir, são explicados alguns destes usos:

- **Setores de Atendimento a Clientes e Call Centers:** um AD pode resolver questões dos usuários sem interação de humanos ou atendentes reais de plantão diretamente no site; se forem usados por atendentes da empresa, agilizam a busca de informações por receberem perguntas em forma natural; neste caso ainda, a forma de armazenamento é mais fácil do que os tradicionais *scripts* e árvores de decisão atualmente utilizados;
- **Intranets:** dúvidas sobre manuais de qualidade e procedimentos técnicos podem ser armazenadas num AD, agilizando a busca por parte de colaboradores;
- **Sites de Comércio Eletrônico:** um AD pode servir para tirar dúvidas sobre produtos e até para fechar uma venda durante a conversa com o usuário;
- **Sites de entretenimento:** um AD sob um personagem virtual pode causar curiosidade atraindo usuários para o site;
- **Treinamento à distância e e-learning:** os AD's podem ser "abastecidos" com conteúdos para dar suporte a instrutores e alunos; imagens podem ser utilizadas no AD para demonstrações práticas;
- **Sistemas de CRM:** informações sobre clientes podem ser coletadas através de um AD e armazenadas em bases de dados, podendo gerar relatórios de pesquisas, ou mesmo para oferecer informações personalizadas ao cliente;
- **Sites em geral para merchandising:** AD's podem ser utilizados para ofertar produtos e serviços para usuários Web;
- **Departamentos de suporte e manutenção:** empresas, principalmente de informática (distribuidores de software e provedores), têm elevados custos com atendimento a chamados (por telefone ou e-mail), sendo que muitas vezes as questões são básicas; um AD pode dirimir dúvidas de clientes antes que um atendente humano precise ser acionado.

Proposta de Assistente Digital

O assistente digital (AD) a ser apresentado neste artigo recebe como entrada uma pergunta de forma textual e escrita. Após, o AD analisa todas as palavras presentes, retirando as chamadas *stopwords*, que são termos muito genéricos sem significado específico (tais como preposições, artigos, etc.). Seguindo, cada palavra é reduzida a seu radical através de um algoritmo de *stemming*, definido especificamente para a língua portuguesa. Para responder à pergunta, o AD procura perguntas semelhantes armazenadas numa base de conhecimento na forma de FAQ's. Para tanto, é necessário que especialistas humanos carreguem esta base com perguntas e respostas. Uma função de similaridade avalia qual das perguntas da base é a mais semelhante à da entrada.

Uma vantagem da implementação proposta é que este especialista não precisa definir exatamente a pergunta, mas somente palavras-chave que, se utilizadas numa pergunta de

um usuário, devem levar à resposta correspondente. Por exemplo, se for cadastrada a pergunta “*cotação dólar 27 maio*”, as perguntas “*podes me dizer a cotação do dólar em 27 de maio*”, “*qual foi a cotação do dólar em 27 de maio*” e “*como ficou a cotação do dólar em 27 de maio*”) levarão à mesma resposta.

Isto segue a premissa de que usuários tendem a utilizar frases curtas, com poucas palavras. Um estudo discutido em (ZUE, 1997) concluiu que clientes de uma organização, interagindo via telefone, utilizam frases curtas (12 palavras ou menos). Neste mesmo estudo, ficou constatado que 70% das frases continham palavras do vocabulário, 11% das frases continham palavras que não estavam no vocabulário (e portanto, geravam confusão de respostas), 11% das frases não tinham palavras significativas (risos ou silêncio), 5% eram frases cortadas e 3% das frases tinham palavras mal pronunciadas que portanto não foram entendidas.

Além disto, o AD proposto neste trabalho também faz tratamento de sinônimos (ex: tu, você, vc), permitindo maior liberdade ao usuário e também diminuindo a possibilidade de erros de interpretação.

A figura 1 apresenta um exemplo de pergunta de um usuário e a respectiva resposta fornecida pelo AD (nesta tela, aparece o AD como sendo utilizado no domínio descrito no estudo de caso).

Armazenamento do Conhecimento

Para responder às perguntas, o AD consulta uma base de conhecimento. Esta base é diferente de um banco de dados (BD's), uma vez no AD o conhecimento é armazenado de forma não estruturada, enquanto que nos BD's existem campos descrevendo atributos de entidades. No AD, há um módulo para cadastramento dos conhecimentos. Estes conhecimentos são cadastrados na forma de perguntas e correspondentes respostas. As perguntas não necessariamente precisam ser cadastradas de forma exata, isto é, partes de perguntas podem ser cadastradas para permitir que pequenas variações sejam atendidas também e de forma correta (conforme já comentado anteriormente).

A base de conhecimento deve ser criada por um especialista do domínio ou por alguém que conheça bem o assunto tratado. Uma lista de FAQ's pode servir de início. Esta base de conhecimento pode ser criada de forma incremental. Uma possibilidade seria utilizar registros passados de soluções de problemas. Uma pessoa analisa as dúvidas possíveis referentes a este conhecimento e gera perguntas que levarão a esta resposta. Desta forma, as próprias pessoas da organização, ao responderem a perguntas de pessoas internas ou de clientes, podem gerar um novo conhecimento para ser armazenado no AD.

Entretanto, é bom que um especialista ou administrador filtre antes estes conhecimentos, para evitar duplicidade, cadastramento de forma errada (tem que conhecer o método) e o cadastro de conhecimento errado (informações incorretas). Por isto, uma lista de FAQ's pode ser montada e, a partir dela, o especialista gera os conhecimentos a serem armazenados no AD.

A figura 2 mostra parte da tela de cadastramento de conhecimentos. Existem basicamente 2 campos básicos na tela da figura 2, que são:

- **Cadastro de perguntas:** aqui a pergunta digitada passa pelas funções de remoção de *stopwords* e de *stemming*; com isso será guardado apenas o radical de cada palavra;
- **Cadastro de respostas:** nesse ponto, a resposta digitada não sofre nenhuma alteração, pois se faz necessário dela para a continuação da conversação. Neste campo optou-se por se cadastrar as respostas no padrão HTML, com isso o especialista na hora do cadastramento de respostas, pode formatar a resposta, inclusive incluindo *links* para páginas Web e *links* para envio de mensagens por e-mail (também podem ser incluídas imagens nas respostas).

Além destes campos, há a possibilidade de se associar uma imagem à cada resposta. Esta imagem pode ser uma figura humana (para, por exemplo, passar uma emoção), uma página Web (por exemplo, onde ir ou encontrar algo) ou a imagem de um produto (demonstrando como usar ou suas características visuais).

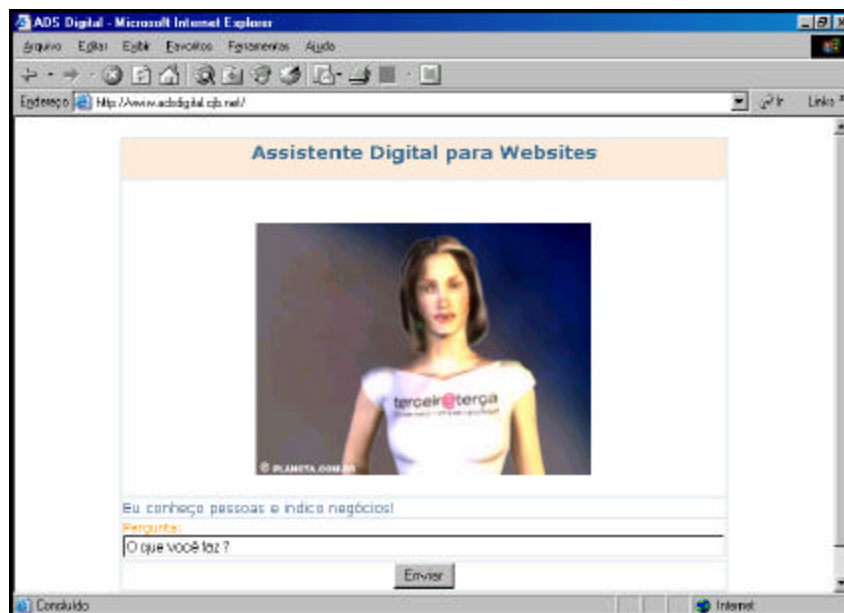


Figura 1: Exemplo de interação com o Assistente Digital (AD)

Outras Funções do AD

O AD ainda possui outras funções para administração da base de conhecimento. Por exemplo, existem funções para inserir, editar e excluir conhecimentos. Há também uma função para inclusão de respostas padrão, que serão exibidas quando a pergunta do usuário não levar a nenhuma resposta.

Também existem relatórios que permitem analisar o comportamento dos usuários interagindo com o AD. Todas as perguntas que não geraram respostas são armazenadas para análise posterior. Isto permitirá aumentar a base de conhecimento, a partir conhecimentos que não se encontram na base. Relatórios estatísticos também permitem descobrir que palavras estão mais sendo usadas pelos usuários. Tal informação é útil para se saber os temas mais procurados.

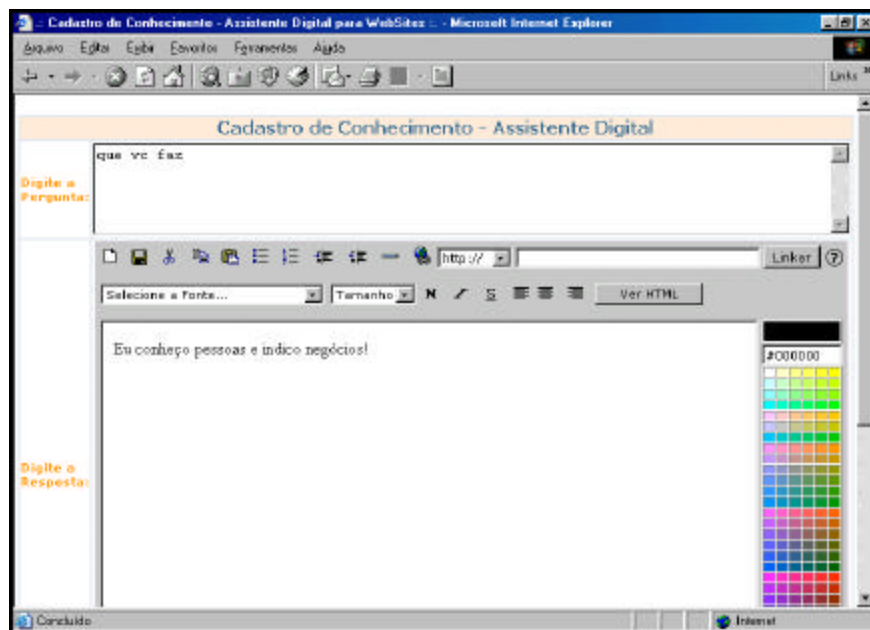


Figura 2: Tela de cadastro de conhecimento no AD proposto

Implementação e Tecnologias Empregadas

O AD foi implementado com tecnologias baseadas em software livre, tais como PHP, Javascript e MySQL.

JavaScript é uma linguagem de programação que proporciona maior interatividade com os usuários. Os programas feitos em JavaScript permitem que praticamente todas as aplicações sejam manipuladas no *browser* do usuário, desde validação de formulários, apresentação de novas janelas, manipulação de imagens, criação de camadas, cálculos complexos, e muitas outras ações que podem tornar as páginas extremamente interativas (SOARES, 2000). Atualmente, a maioria dos *browsers* suporta JavaScript.

Já PHP é uma linguagem de script para uso em servidores, e o seu script é processado pelo servidor Web. O servidor faz a codificação dos códigos PHP, retornando para o navegador. Este tipo de interação permite algumas operações complexas, assim como maior segurança, na hora de leitura de formulários.

O PHP é uma linguagem de criação que permite a elaboração de páginas dinâmicas. Seu código é inserido dentro do corpo HTML, funcionando como *scripts* de linguagens. Sua diferença básica em relação a outros *scripts*, tais com CGI, é a inexistência de múltiplos comandos que outras linguagens exigem junto ao HTML. Os códigos de PHP são escritos dentro de *tags* especiais, que permitem a ativação e desativação do "modo PHP".

O MySQL segundo CHOI et al. (2001) é um sistema de gerência de bancos de dados relacional (SGBDR) que juntou-se à *Open Source Community* apenas recentemente, quando foi lançado pela GNU *Public License*. Além de estar disponível gratuitamente, o MySQL fornece uma enorme quantidade de outros recursos:

- *Software de fonte aberta:* o pacote do MySQL vem com o código-fonte completo. Isso significa que se pode estudar o código de origem e modificá-lo para adequar-se a necessidades particulares;
- *Suporte a SQL:* o MySQL suporta a SQL (*Structured Query Language*), uma linguagem padrão usada para fazer consultas a bancos de dados relacionais. Embora a sintaxe da SQL seja um pouco diferente de um SGBDR para outro, os conceitos subjacentes são praticamente os mesmos. O MySQL usa o seu próprio conjunto de comandos SQL, mas o conceito básico ainda permanece;
- *Performance e confiabilidade:* o MySQL é notavelmente rápido e confiável em um ambiente exigente;
- *Facilidade de uso:* o MySQL é um sistema gerenciador de banco de dados relativamente simples, apesar de seus recursos poderosos. De fato, os bancos de dados do MySQL não são nada mais do que um conjunto de diretórios em que tabelas de bancos de dados são mantidas em arquivos separados. Remover um banco de dados ou uma tabela de banco de dados é, na verdade, simplesmente excluir um diretório ou um arquivo. O MySQL também é simples de ser instalado e gerenciado. Ele vem com um conjunto completo de aplicativos cliente/servidor além de vários utilitários que facilitam a administração do banco de dados;
- *Suporte gratuito:* pode-se conseguir ajuda de outros usuários que utilizam o MySQL por vários meios: o grupo de notícias MySQL, grupos de discussão, sites Web independentes que compartilham gratuitamente o seu conhecimento, entre outros. Precisando de suporte técnico de nível mais alto, pode-se consegui-lo com projetistas e consultores a um custo mínimo;
- *Troca de plataformas:* o MySQL pode ser executado tanto em plataformas UNIX como Windows.

Técnicas Empregadas

Na literatura, o problema da variação morfológica das palavras é frequentemente denominado *stemming*. A palavra *stem* vem do inglês e pode ser entendida como origem, raiz. Também se denomina *stemming* todas aquelas práticas que buscam tratar o problema. Alguns mecanismos de procura na *Web*, como o *Altavista*, utilizam a idéia do *stemming* para qualificar a consulta do usuário pesquisando também pelas variantes das palavras fornecidas como critério.

Existe uma certa aceitação de que a remoção de sufixos, reduzindo palavras semelhantes a uma forma padrão, é uma boa técnica. Em um processo de indexação de documentos, por exemplo, a técnica de *stemming* reduziria o tamanho do vocabulário dos documentos, já que palavras de mesma origem seriam mapeadas a um formato padrão. Porém, também se discute a real contribuição que a técnica efetivamente traria, ou até possíveis perdas de representação já que algumas variações morfológicas são algumas vezes necessárias para representar diferentes conceitos.

Para ser utilizado no presente trabalho, foram investigadas as soluções existentes de *stemmers* para a língua portuguesa. A grande maioria dos algoritmos trata a variação morfológica na língua inglesa, e os trabalhos realizados para o português ainda são muito poucos.

Com base nesses requisitos, foi escolhida uma versão do algoritmo de Paice/Husk para o português, desenvolvida por ORENGO & HUYCK (2001). Este algoritmo, e mais especificamente esta versão, possui grande flexibilidade na definição de regras e no tratamento de exceções, além do seu funcionamento ser de fácil compreensão. Além disso, o trabalho de ORENGO & HUYCK (2001) representa uma importante fonte de dados coletados a partir de experiências com algoritmos de *stemming* aplicados na língua portuguesa.

Estudo de Caso

Um dos casos onde o AD está sendo utilizado é no Portal do Terceira Terça (www.terceiraterca.com.br), uma ONG responsável por eventos de *networking* (contatos entre empresas e profissionais). O Terceira Terça foi concebido como um evento voltado a empresas e profissionais da área de TI, com objetivos de proporcionar a divulgação das empresas (idealizadores, patrocinadores, apoiadores, parceiros e empresas participantes), promover a valorização dos profissionais do mercado e fomentar a geração de negócios entre os participantes. O evento é mensal, acontecendo toda terceira terça-feira do mês. Ele é realizado em diversas cidades: Porto Alegre, Caxias do Sul, Novo Hamburgo, Florianópolis e em breve estará em São Paulo.

Cerca de 500 pessoas em média são recebidas a cada edição e atualmente (após 23 edições) são mais de 700 empresas e 10.000 usuários cadastrados no portal. Os contatos feitos no evento já geraram negócios da ordem de mais de 2 milhões de reais, conforme estimado pelos organizadores a partir de relatos das empresas participantes.

Neste portal, é utilizada a musa virtual do Terceira Terça (a Tetê), a qual será a personagem do assistente digital. No caso da Tetê, o AD tem por finalidade fornecer informações sobre o evento, tais como onde acontece, quando, para que serve, como se inscrever, como chegar até lá, etc. Este AD pode ser experimentando a partir do portal da Terceira Terça (www.terceiraterca.com.br/ad) ou do site www.adsdigital.com.br/adsdigital.

No momento, existem quase 500 conhecimentos armazenados (perguntas e respostas). Alguns destes são referentes ao domínio (informações sobre o evento), outros são sobre cultura em geral (Bush e Bin Laden, por exemplo), outros definem a personalidade da personagem (por exemplo, gostos) e outros ainda existem apenas para manter a conversação.

Também é função do AD apresentar pessoas e empresas previamente cadastradas, bem como indicar empresas (usuário pode solicitar lista de empresas de um certo ramo ou que trabalhem em certos tipos de negócios). Estas informações específicas são fornecidas de forma dinâmica, através de consultas em um banco de dados estruturado auxiliar, onde constam nomes de empresas, endereços, tipo de negócios, etc. Assim, o AD pode gerar novos tipos de respostas, mesmo que com um número limitado de FAQ's, pois as respostas são geradas dinamicamente com dados desta base auxiliar.

Neste estudo de caso, o AD possui uma identidade visual e uma personalidade, simulando uma atendente ou *promoteur* virtual, chamada de Tetê. Graças ao auxílio da empresa Plancta (www.plancta.com.br), especializada em design de personagens digitais, o AD do Terceira Terça possui uma imagem digital (de uma modelo criada artificialmente) e pode ter emoções, já que diferentes feições e gestos podem ser apresentados, dependendo do tipo de resposta fornecida.

Entretanto, o AD em geral não precisa ter um modelo digital. É possível utilizar fotos de pessoas, produtos ou páginas do próprio site como imagens associadas às respostas.

Por fim, o AD também está sendo utilizado para *merchandising*, podendo divulgar empresas e marcas e mesmo fazer ofertas de produtos e serviços. Tais oferecimentos são acionados a partir de determinadas perguntas do usuário (com conhecimentos específicos previamente cadastrados na base) ou podem surgir por inatividade do usuário.

Conclusões

Este trabalho pretendeu demonstrar a importância e a viabilidade de assistentes digitais (AD's) para auxiliar usuários em Portais ou sites Web. A principal vantagem está em diminuir o tempo e esforço para armazenar e recuperar conhecimento, uma vez que o conhecimento é formatado de forma similar a FAQ's (*Frequently Asked Questions*). Para se ter um parâmetro, no caso da Tetê (o AD da Terceira Terça), os 300 primeiros conhecimentos foram armazenados em aproximadamente 3 horas (dispersas em 2 dias).

O emprego de AD's pode melhorar a Gestão do Conhecimento nas organizações. O exemplo do AD descrito no estudo de caso confirma que os usuários encontram mais rápida e facilmente as informações do que navegando num Web site. Além disto, o AD conquista a simpatia dos usuários, uma vez que interage em linguagem natural e também por possuir um "rosto" amigável, o que contribui para divulgar o serviço. O AD ainda torna a conversação mais pessoal, pois passa ao usuário a sensação agradável de se estar conversando com uma pessoa.

AD's como o proposto aqui podem beneficiar diversos tipos de serviços como Call Centers, Intranets, Setores de Atendimento ao Cliente, departamentos de suporte, ouvidorias, programas de qualidade, sistemas de CRM e até mesmo pode ser utilizado para *merchandising*, fornecendo informações e imagens de produtos e serviços oferecidos.

Algumas limitações ainda existem no AD proposto. Nem todas as perguntas irão gerar respostas adequadas. Isto porque a tecnologia de Inteligência Artificial existente ainda não consegue o mesmo desempenho de um interlocutor humano. Entretanto, para o fim estabelecido, que é fornecer informações sobre um domínio específico (e limitado), o AD mantém um grau de aceitação alto.

Apesar da eficácia e eficiência da tecnologia empregada, o AD não será útil se a base de conhecimento não for de qualidade. Isto quer dizer que, primeiro, somente conhecimentos corretos devem ser armazenados, e isto deve ser feito por alguém que conheça o assunto do domínio e que saiba como, didaticamente e de forma escrita, passar este conhecimento para outros. No AD proposto, somente administradores cadastrados podem armazenar conhecimento.

Entretanto, todas as pessoas podem contribuir para o armazenamento de conhecimento fornecendo perguntas e respostas como se fosse uma lista de FAQ's, para que os administradores possam aumentar a base de conhecimento. Assim, é importante a conscientização de cada um para compartilhar seu conhecimento com os outros. Se isto não ocorrer, de nada adiantará a moderna tecnologia empregada no AD.

Por fim, uma característica importante deste AD é a sua generalidade, ou seja, dependendo do contexto onde ele será empregado, a sua função muda, bem como o conhecimento nele contido. Para tanto, o AD foi desenvolvido de forma a facilitar o armazenamento de novos conhecimentos na base.

Trabalhos futuros

Inúmeros serviços e aperfeiçoamentos estão previstos para o AD apresentado neste trabalho. Uma das propostas é o AD possuir voz (sintetização de voz a partir das respostas).

Também estão em desenvolvimento os seguintes melhoramentos:

- a) análise do contexto: hoje as perguntas são independentes; o objetivo futuro é manter o contexto de uma pergunta para outra;
- b) consultas em bancos de dados próprios da empresa: está em teste um serviço de busca de dados em bancos existentes; no caso do Terceira Terça, o AD poderá encontrar empresas a partir de ramos fornecidos pelo usuário, bem como endereços e sites de empresas a partir de nomes fornecidos;
- c) coleta de dados: o AD fará perguntas e armazenará as respostas do usuário em bancos de dados da organização; por exemplo, para pesquisas, inscrições ou cadastramentos, vendas e pedidos pelo AD;
- d) funções básicas tais como fornecimento de dia e hora.

O AD também deverá em breve conter funções para *merchandising*, podendo oferecer produtos e serviços, ou divulgar empresas patrocinadoras. Da mesma forma, planeja-se acrescentar no AD um serviço de notícias vindas de um site específico (www.baguete.com.br). Assim, o AD poderá ser enriquecido com informações sobre determinados assuntos e de forma automática, sem a necessidade de cadastro por pessoas. Um sistema deste tipo foi apresentado em (SALDANÃ et al., 2003), onde informações específicas como a cotação do dólar e previsão do tempo são armazenadas automaticamente sem intervenção humana. Atualmente, a cotação do dólar está em fase de teste no AD para a Terceira Terça.

Por fim, mais tarde, espera-se também poder fazer o reconhecimento de voz das perguntas para que os usuários possam utilizar os serviços do AD por telefone. Isto será extremamente útil para portais de voz e para atendimento automático em Call Centers.

Referências Bibliográficas

- CHEN, Hsinchum (1994) The vocabulary problem in collaboration. **Computer**, 27(5), p.2-10, May.
- CHOI, Wankyu; KENT, Allan; PRASAD, Ganesh; ULLMAN, Chris. (2001) **Beginning PHP4 Programando**. São Paulo: Makron Books, 719p.
- DAVENPORT, T. H & PRUZAC, L. (1997) **Working knowledge – how organizations manage what they know**. Harvard Business School Press, 224 p.
- ORENGO, Viviane M & HUYCK, Christian. (2001) A Stemming Algorithm for the Portuguese Language. In: Symposium on String Processing and Information Retrieval, SPIRE'2001, Chile.
- SALDAÑA, Ramiro et al. (2003) Captura automática de textos en la web para bibliotecas digitales. In: IV Coloquio Internacional de Ciencias de la Documentación y VI Congreso del Capítulo Español de ISKO (International Society for Knowledge Organization), Salamanca, Maio.
- SENGE, Peter M. (2001) **A quinta disciplina: arte e prática da organização que aprende**. 9. ed. Best Seller, 444p.
- SOARES, Wallace (2000) Programando em PHP Conceitos e Aplicações. São Paulo: Editora Érica. 2. ed. 400p.
- WEIZENBAUM, J. (1967) ELIZA: a computer program for the study of natural language communication between man and machine. **Communications of the ACM**, v.10.
- ZUE, Victor (1997) Conversational interfaces: advances and challenges. In: **Proceedings EUROSPEECH'97**. Rhodes, Grécia, p.9-18.