

Um Ambiente de e-Learning Com Suporte a Reconhecimento de Voz

Leonardo O. Moreira¹, Jonas A. Nascimento¹, José R. O. Damico²,
Nelson E. N. Oliveira¹, Javam C. Machado¹

¹ Departamento de Ciência da Computação (DC)
Universidade Federal do Ceará (UFC) – Fortaleza, CE – Brasil

²IBM Máquinas Industriais e Serviços
Rodovia SP-101 – 13.175-000 – Hortolândia – SP - Brasil

leoomoreira@ufc.br, jonas.araujo@gmail.com, jdamico@br.ibm.com,
nelsoneldoro@gmail.com, javam@ufc.br

Abstract. *e-Learning environments have as purpose to provide distance learning using technologies which support that benefit. English is widely adopted around the world, and the global market has been requiring the employees have, at least, basic English skills. There are many e-learning environments, but none of them provides a voice recognition mechanism, essential for foreign language learning. This article presents iVeLA, open-source e-learning environment built over an extensible architecture and with voice recognition. Experiments were developed and analyzed to measure the accuracy of the voice recognition strategy. The results of those experiments proved that the strategy could be used on iVeLA.*

Resumo. *Ambientes de e-Learning tem o propósito de prover ensino à distância mediante do uso de tecnologias que concebem este benefício. O inglês é o idioma adotado como universal, e o mercado empregatício global vem exigindo no mínimo conhecimentos básicos das pessoas neste idioma. Existem diversos ambientes de e-Learning, mas nenhum deles possui recursos de reconhecimento de voz, indispensável para o ensino de línguas estrangeiras. Este artigo apresenta o iVeLA como um ambiente de e-Learning de código aberto, dotado de uma arquitetura extensível e possuindo recursos de reconhecimento de voz. Experimentos foram feitos e analisados para avaliar a eficácia da estratégia de reconhecimento de voz do iVeLA.*

1. Introdução

Em diferentes momentos históricos, desde o nascimento da escola, diferentes metodologias de ensino foram sendo utilizadas com o objetivo de melhorar sua qualidade. A metodologia presencial é a modalidade adotada na maioria das instituições de todo o mundo. Com a popularização e o crescimento da Web, ambientes de *e-learning* estão sendo desenvolvidos no intuito de prover meios para o ensino à distância. O conhecimento da língua inglesa é um dos requisitos básicos para as pessoas ingressarem no mercado empregatício global. Nesse sentido, é importante haver uma certa facilidade de acesso ao ensino deste idioma. Os ambientes de *e-learning* fornecem recursos de acesso ao conhecimento, mas

a maioria dos ambientes de ensino à distância, existentes, é ineficiente para o ensino de línguas estrangeiras.

Este artigo descreve o desenvolvimento de um ambiente para ensino à distância com reconhecimento de voz voltado para o ensino da língua inglesa, denominado iVeLA (Internet Voice e-Learning Application), que tem como objetivo disseminar educação assistida por computador em ambiente Web, capaz de reconhecer e analisar a pronúncia dos usuários por meio de um módulo para reconhecimento de voz. Este ambiente de ensino/aprendizagem disponibiliza uma interface para funcionalidades convencionais de educação à distância, como salas de bate-papo, fóruns de discussão, *download* e *upload* de arquivos, sistemas de avaliações, agenda e sistemas de comunicação entre usuários.

O iVeLA foi validado como aplicação de educação à distância assistida por computador, contendo um curso de inglês para negócios. A interface do sistema para interação com o usuário foi projetada e implementada de forma a facilitar a utilização da ferramenta por parte dos usuários, administradores, professores, tutores e alunos. Para isso, foram utilizados recursos da Web 2.0, que tornam as interfaces Web mais poderosas, elegantes, fáceis de utilizar e de disponibilizar informações no ambiente. O iVeLA foi desenvolvido pela parceria entre UFC [UFC 2009] e IBM [IBM 2009] com financiamento da Flextronics [Flextronics 2009].

2. Trabalhos Relacionados

Blackboard [Wainwright et al. 2007] é um ambiente de ensino e aprendizado considerado líder em plataforma de *e-learning* em virtude de sua facilidade de uso, larga adoção em instituições de ensino, flexibilidade pedagógica, amplitude de funções e características intuitivas, para aprendizado à distância e apoio ao ensino presencial. Ele oferece ambiente amigável para gerenciamento de cursos, desenvolvimento de conteúdo, avaliações e promoção de atividades de colaboração e trabalhos desenvolvidos em grupo.

O ambiente WebCT (Web Course Tools) [Fuller et al. 2001] fornece várias ferramentas para a criação de cursos baseados na Web, podendo ser usado para criar cursos *on-line* completos, ou como interface de apoio para cursos comuns. Este ambiente tem como principais características: interface amigável, não exigência de conhecimentos específicos sobre Internet por parte dos professores e dos alunos, disponibilidade de ferramentas que facilitem a aprendizagem, a comunicação e a colaboração, além de facilitar a administração do curso.

Moodle [Cole and Foster 2007] [Wainwright et al. 2007] é um *software* livre, de apoio à aprendizagem, executado em ambiente virtual. Este consiste em sistema de administração de atividades educacionais destinado à criação de comunidades *on-line*, em ambientes virtuais voltados para a aprendizagem colaborativa. O Moodle é o único sistema *open-source*, atualmente disponível, que pode competir com os grandes sistemas comerciais.

Os ambientes descritos anteriormente contemplam os requisitos básicos necessários para disseminar ensino à distância, mas, por não possuírem recursos de reconhecimento de voz, são ineficientes para o ensino à distância de línguas estrangeiras. Além disso, esses ambientes não atendem aos requisitos de uma arquitetura orientada a serviços, necessários para prover a extensibilidade, com qualidade, no aplicativo.

3. iVeLA

O iVeLA é um aplicativo para ensino à distância em ambiente Web com reconhecimento de voz. Este ambiente Web de apoio ao processo ensino/aprendizagem disponibiliza a interface para funcionalidades convencionais básicas de educação à distância, como bate-papo, fóruns de discussão, sistemas de comunicação, exercícios e avaliações, *download* e *upload* de arquivos.

Ferramentas de bate-papo são de grande importância no provimento de aulas virtuais, avaliação de alunos, etc. A maioria das ferramentas de bate-papo ou salas virtuais, utilizadas por outros ambientes de educação à distância, possui deficiências devido à tecnologia utilizada para prover este serviço na Web. O iVeLA implementa um serviço de salas de bate-papo robusto e escalável baseado no IRC (Internet Relay Chat) [Charalabidis 2000] que é um serviço bastante disseminado e conhecido. O serviço IRC, existente no iVeLA, foi escrito em Java utilizando a biblioteca de código aberto chamada JerkLib [JerkLib 2009]. Um diferencial na ferramenta de bate-papo do iVeLA é a função blackboard. O blackboard simula um quadro negro virtual onde o professor, responsável pela sala de bate-papo, pode desenhar rascunhos que são visualizados remotamente pelos alunos.

A extensibilidade constitui um dos principais requisitos do iVeLA; para tanto, o sistema tem uma arquitetura orientada a serviços, facilitando a inclusão de novos serviços produzidos por terceiros, bem como fornecendo serviços para sistemas com arquitetura semelhante. Através dessa arquitetura, novas funcionalidades poderão ser disponibilizadas por meio da interoperabilidade de serviços do iVeLA com outros ambientes de ensino e aprendizagem. O ambiente iVeLA está disponibilizado para a comunidade, como um sistema de *software* livre, código aberto (SL/CA), sob uma licença do tipo GPL [GNU 2009]. Desta forma, provedores de conteúdo poderão fazer uso do aplicativo para disponibilizar quaisquer cursos, em caráter aberto, para a comunidade em geral.

3.1. Arquitetura

O iVeLA é composto por quatro grandes camadas. Nesta seção, serão descritas essas quatro camadas, seus objetivos e as tecnologias contidas em cada uma. A Figura 1 ilustra uma visão de tecnologias e camadas da arquitetura do iVeLA.

A camada mais próxima do usuário, a camada de visão, promove a interação do usuário com o ambiente. Esta camada utiliza recursos da Web 2.0, fornecidos pelas bibliotecas AJAX [Dwyer 2008] contidas na arquitetura. As páginas Web são escritas utilizando a tecnologia JSP (Java Server Pages) [Sun 2009], que fornece o dinamismo contido nas páginas do ambiente iVeLA. Outro recurso importante da camada de visão é a utilização do SiteMesh *framework* [Dwyer 2008] que fornece uma modularização, o que força o ambiente a ter um padrão de estrutura de visualização. Além disso, o SiteMesh promove interfaces de fácil modificação, vez que são modularizadas.

A camada de controle é responsável por orquestrar o fluxo do processo do aplicativo, ou seja, receber as requisições do usuário, processar ou delegar para a camada de negócio e devolver o resultado do processamento. Além disso, a camada de controle contém funções de serviços para prover a interoperabilidade no iVeLA; com isso, outras aplicações ou ambientes poderão fazer uso de recursos existentes do sistema. Para a implementação de serviços Web a tecnologia REST [Shi 2006] é utilizada. Já para o

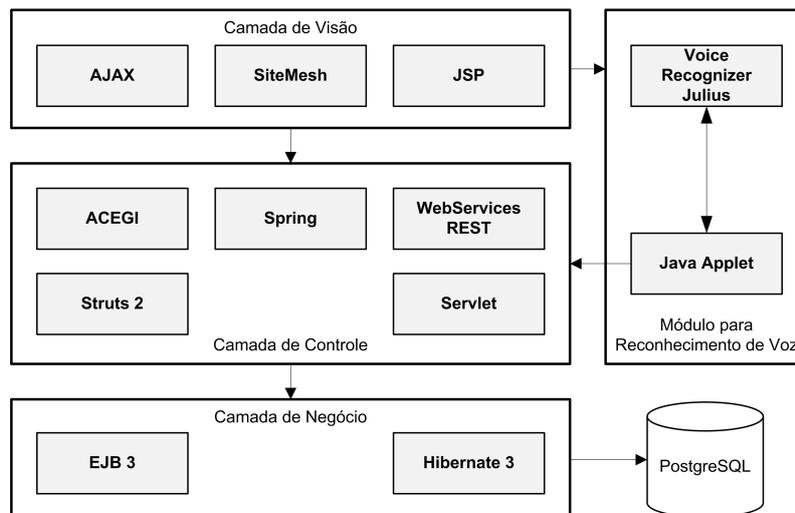


Figura 1. Arquitetura do iVeLA

serviço de autenticação e autorização dos usuários no iVeLA, o ACEGI [Dwyer 2008] é integrado ao ambiente. O ACEGI é um componente de *software* livre gerenciado pelo Spring *framework* [Velde et al. 2007] e provê este tipo de serviço para aplicações corporativas escritas em Java.

O Struts 2 *framework* [Brown et al. 2008] é um poderoso componente para forçar a utilização do padrão MVC (Model-View-Controller) e facilitar o desenvolvimento de aplicações corporativas. Ele é utilizado para a orquestração de fluxos de processos e de chamadas à camada de negócio ou de serviços. O Spring também é um componente de *software* livre, baseado em padrões de projeto, que fornece injeção de dependências para aplicações escritas em Java. A injeção de dependências é extremamente importante para fornecer um baixo acoplamento entre classes de uma aplicação orientada a objetos. O Spring, no iVeLA, é utilizado para injetar objetos em classes que possuem determinadas características, e assim facilitar o desenvolvimento e a integração entre os diversos *frameworks* e tecnologias existentes na arquitetura. Os Servlets [Sun 2009] são componentes da plataforma J2EE (Java 2 Enterprise Edition) do lado servidor, baseados no modelo requisição/resposta para realizar algum processamento. Os Servlets são utilizados no iVeLA para a implementação de recursos que não se encaixam como serviços Web ou na camada de controle.

Já a camada de negócio é composta pela tecnologia EJB (Enterprise Java Bean) 3.0 [Sun 2009] e pela camada de persistência Hibernate 3.0 [Bauer and King 2006]. Os EJBs são componentes que fazem parte da plataforma J2EE voltados para a implementação da camada de negócio das aplicações. A utilização de componentes EJBs torna as aplicações menos acopladas e possibilita a reutilização dos componentes de negócios em outras aplicações. O Hibernate 3.0 é uma implementação da JPA (Java Persistence API) utilizada para prover a manipulação e persistência de dados. Ele fornece a abstração de dados no modelo objeto, tornando assim mais fácil a manipulação dos dados em linguagem orientada a objetos como Java. A última camada, a de reconhecimento de voz, é responsável por captar sinais de voz do usuário, tratá-los, e realizar alguma computação. Esta camada é o diferencial do iVeLA sob o ponto de vista de inovação. Para prover tal

recurso, esta camada se serve do Julius [Julius 2009], que é um motor para reconhecimento de voz. Uma *applet* [Sun 2009] é utilizado para captar os sinais de voz do usuário, enviá-los ao Julius e retornar os resultados através da camada de controle.

3.2. Módulo para o Reconhecimento de Voz

O módulo de reconhecimento de voz é responsável por disponibilizar os exercícios de voz no iVeLA aos alunos de um Curso de Inglês. Para implementar o exercício de voz, usam-se duas tecnologias: *applets* Java e o reconhecedor de voz Julius [Julius 2009]. O *applet* de reconhecimento de voz tem como tarefa implementar uma camada entre o usuário e a *engine* de reconhecimento de voz. Além disso, o *applet* é tem a função de instalar o Julius na máquina do cliente, caso não exista.

A *engine* de reconhecimento de voz funciona da seguinte forma: dada uma gramática, um vocabulário e um modelo acústico, a frase ou palavra do vocabulário é reconhecida pelo *applet*, ao ser pronunciada no microfone. A gramática define um conjunto de regras que são usadas pelo Julius, com o intuito de formar a frase. O vocabulário compreende um conjunto de palavras disponíveis para reconhecimento. O reconhecimento só é possível graças ao modelo acústico. Este é uma amostragem de gravações feitas em ambientes otimizados, com o objetivo de angariar pronúncias de diversos fonemas, formando assim um “banco de dados” fonético. O modelo acústico é disponibilizado pelo grupo americano VoxForge [VoxForge 2009], como *software* livre.

4. Experimentos

Como mostrado anteriormente, o iVeLA contempla as principais funcionalidades existentes na maioria dos ambientes para educação à distância. O principal diferencial do iVeLA está no recurso de reconhecimento de voz. Para testar a eficácia e escalabilidade do *applet* e do reconhecedor de voz, fizemos testes com o sistema, usando diversos alunos em diferentes níveis de inglês. Os testes consistem em cinco frases em inglês: (1) “My name is Bob”, (2) “I am from Canada”, (3) “Here is my card”, (4) “He is Jack” e (5) “He is French”. No experimento, essas frases devem ser escutadas e repetidas pelo usuário. Todas as cinco frases somadas foram repetidas 930 vezes, ou seja, cada frase foi repetida 186 vezes. O aplicativo avaliou os alunos da seguinte forma: (A) alunos que acertaram a frase na primeira tentativa; (B) alunos que não acertaram a frase de forma alguma; (C) alunos que acertaram uma mesma frase a partir da segunda tentativa. A Tabela 1 resume o resultado para cada frase.

Tabela 1. Resultados do experimento

	Frase 1	Frase 2	Frase 3	Frase 4	Frase 5	Total
A	128	146	119	129	142	664
B	3	7	32	9	6	57
C	55	33	35	53	33	209
Total	186	186	186	186	186	930

Os testes de campo mostram que apenas 57 amostras não foram, de forma alguma, reconhecidas pelo *applet*. Isso se deve ao fato de que o modelo acústico disponível ainda está em fase experimental, além de não ser adequado a todos os tipos de entonações possíveis. Do total de acertos, podemos ver que 873 amostras (664 + 209), de um total de 930, foram faladas e reconhecidas de forma correta, mesmo após a segunda tentativa. Isso mostra que, tanto o *applet* quanto o reconhecedor de voz, obtiveram uma boa aceitação.

5. Conclusão

Neste artigo foram demonstrados o iVeLA, suas funcionalidades, sua arquitetura e as inovações. O iVeLA contempla todas as funcionalidades desejadas para prover um ambiente eficaz para o ensino à distância. Além disso, o iVeLA utiliza recursos da Web 2.0 que torna sua interface mais atrativa e usual. Muitos ambientes de educação à distância possuem deficiências em suas ferramentas de bate-papo. O iVeLA se serve da tecnologia IRC para prover uma ferramenta de bate-papo mais robusta.

O iVeLA possui uma arquitetura modular, flexível e escalável, que o torna mais robusto, com mais disponibilidade e extensibilidade. Além disso, os recursos de serviços Web, contidos no iVeLA, tornam este último interoperável com outros ambientes. Experimentos realizados com o módulo de reconhecimento de voz foram feitos e seus resultados analisados. Constatou-se a eficácia de modo aceitável para este recurso. Assim, o iVeLA é a ferramenta que mais se adequa para o ensino da língua inglesa dentre os outros ambientes analisados.

Referências

- Bauer, C. and King, G. (2006). *Java Persistence with Hibernate*. Manning Publications Co., Greenwich, CT, USA.
- Brown, D., Davis, C., and Stanlick, S. (2008). *Struts 2 in Action (In Action)*. Manning Publications Co., Greenwich, CT, USA.
- Charalabidis, A. (2000). *The Book of IRC*. No Starch Press, San Francisco, CA, USA.
- Cole, J. and Foster, H. (2007). *Using moodle, 2nd edition*. O'Reilly.
- Dwyer, J. (2008). *Pro Web 2.0 Application Development with GWT (Pro)*. APress.
- Flextronics (2009). *Flextronics Instituto de Tecnologia*. <http://www.fit-tecnologia.org.br/>.
- Fuller, A., Awyzio, G., and McFarlane, P. (2001). Using webct to support team teaching. pages 315–318.
- GNU (2009). *The GNU General Public License*. <http://www.gnu.org/copyleft/gpl.html>.
- IBM (2009). *IBM Brasil*. <http://www.ibm.com/br/pt/>.
- JerkLib (2009). *JerkLib - A Java IRC Library*. <http://jerklib.sourceforge.net/>.
- Julius (2009). *Julius Voice Recognizer Home Page*. <http://julius.sourceforge.jp/>.
- Shi, X. (2006). Sharing service semantics using soap-based and rest web services. *IT Professional*, 8(2):18–24.
- Sun (2009). *Java EE at a Glance*. <http://java.sun.com/javaee/>.
- UFC (2009). *Universidade Federal do Ceará*. <http://www.ufc.br/>.
- Velde, T. V. d., Dupuis, C., Balani, N., and Li, S. (2007). *Beginning Spring Framework 2*. John Wiley & Sons, Inc., New York, NY, USA.
- VoxForge (2009). *VoxForge Acoustic Model*. <http://voxforge.org/home>.
- Wainwright, K., Osterman, M., Finnerman, C., and Hill, B. (2007). Traversing the lms terrain. In *SIGUCCS '07: Proceedings of the 35th annual ACM SIGUCCS conference on User services*, pages 355–359, New York, NY, USA. ACM.