

ESTUDO E APLICAÇÃO DE RECURSOS DE ACESSIBILIDADE NO APLICATIVO “ALERTA BRUSQUE”

Study and application to accessibility resource in application “Alerta Brusque”

Pedro Sidnei Zanchett ¹

Wagner Correia ²

Lucas Debatin ²

Ivan Fumagalli Varela ²

Resumo: O uso de interfaces naturais em sistemas computacionais vem aumentando devido à evolução das tecnologias de toque, gesto e voz, permitindo criar aplicações extremamente ricas especialmente para pessoas com pouca experiência ou com necessidades especiais. Em situações de emergência, como nos apresenta o aplicativo Alerta Brusque, é fundamental a fácil interação dos usuários com o aplicativo para localizar rapidamente as informações de acesso aos níveis de rio e chuva em situação de preocupação extrema. Este artigo refere-se a um projeto de pesquisa que visa implementar funcionalidades de comando por voz e resposta automática por áudio das informações sobre nível de rio e chuva do rio Itajaí Mirim no aplicativo Alerta Brusque, através da conversão da voz em comandos inteligentes e converter sua resposta em voz atendendo de forma imediata às necessidades do usuário, aumentando a relação do usuário com dispositivo móvel.

Palavras-chave: Experiência de usuário. Usabilidade. Reconhecimento de voz. Leitura em áudio.

Abstract: The use of natural interfaces in computer system has been rising due to evolution of touch, gesture and voice technologies, allowing you to create extremely rich applications especially for people with little experience or special needs. In emergency situations such as the present Alerta Brusque application is vital the easy user interaction with the application to find quickly access information on river levels and rainfall in estrema concerns situation. This article refers to research project aims to implement the voice control and automatic answer for audio about the informations of Itajaí Mirim river and rain in Alerta Brusque application by converting voice into smart commands and convert your answer in voice given immediately user needs increasing user relationship with mobile.

Keywords: User experience. Usability. Voice recognition. Audio reading.

Introdução

A região do Vale do Itajaí tem em seu histórico enchentes, alagamentos e deslizamentos. Estes fenômenos naturais causam grandes perdas para a sociedade, como vidas humanas, desabrigados, e também impacta no desenvolvimento econômico da região. Esta realidade fez o acadêmico Wagner Correia do curso de Sistemas de Informação questionar-se sobre como poderíamos auxiliar esta população, já que o uso de equipamentos móveis, como *smartphones* e *tablets* é alto e demonstra crescimento.

Por meio do projeto de pesquisa do Artigo 171 foram lançados, em 2015, dois aplicativos: o “Alerta Brusque” e o “SIAH” (Sistema de Informação de Análises Hidrológicas), aplicativos que apresentam para a população em tempo real os níveis do rio e de chuva do Rio Itajaí Mirim.

Além das informações sobre o rio e chuva, a população brusquense pode, por meio do aplicativo, consultar telefones úteis às épocas de emergência, *sites* de utilidade pública, notícias

¹ Mestre em Engenharia e Gestão do Conhecimento pela Universidade Federal de Santa Catarina (UFSC). Docente no Centro Universitário Leonardo Da Vinci – UNIASSELVI e Centro Universitário de Brusque – UNIFEBE. E-mail: pedrozanchett@gmail.com

² Acadêmico do curso de Sistemas de Informação pelo Centro Universitário de Brusque – UNIFEBE.

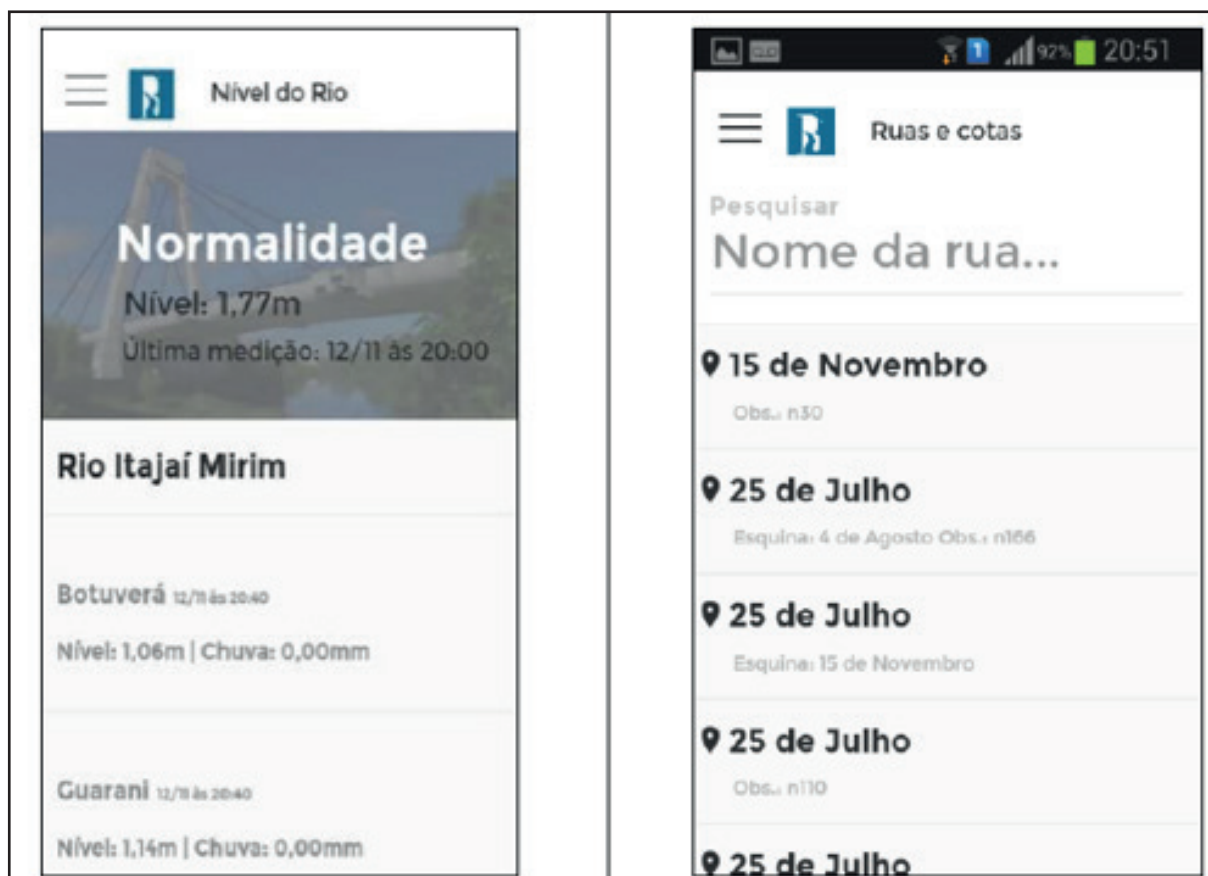
etc. O aplicativo foi desenvolvido em parceria com a Defesa Civil de Brusque e com apoio da Prefeitura Municipal de Brusque.

Atualmente, as fontes de dados utilizadas pelo Alerta Brusque são: Defesa Civil de Brusque, CEOPS (Centro de Operação do Sistema de Alerta - FURB) e ANA (Agência Nacional das Águas).

O Alerta Brusque é integrado ao SIAH por meio de uma API (*Application Programming Interface*), com saída de dados em formato JSON para Android e iOS. O aplicativo oferece à população acesso às informações geradas pelas estações automatizadas dos principais pontos da cidade de Brusque, sendo possível consultar informações que incidem sobre alguns bairros de Brusque, onde cada estação de telemetria que a cidade possui realiza a transmissão de dados em períodos regulares de 10 minutos.

A versão Android está disponível no *link* <<https://goo.gl/Yn8PSD>> e para iOS em <<https://goo.gl/T1LNdy>>. O aplicativo Alerta Brusque é apresentado na Figura 1, onde está a tela inicial na versão para iPhone e a tela de Ruas e Cotas num *smartphone* com Android.

Figura 1. Telas do Aplicativo Alerta Brusque no Android e iOS



Fonte: Elaborado pelos autores (2016)

O Alerta Brusque já é uma ferramenta bem conhecida na região, pois durante seu lançamento, em novembro de 2015, houve muitas reportagens em jornais, TVs, rádios e revistas conforme a reportagem concedida à emissora RBS-TV: “Estudante cria aplicativo para ajudar moradores de Brusque em situação de enchente” (Disponível em: <<http://g1.globo.com/sc/santa-catarina/jornal-do-almoco/videos/t/blumenau/v/estudante-cria-aplicativo-para-ajudar-moradores-de-brusque-em-situacao-de-enchente/4617904/>>. Acesso em: 3 jun. 2016).

Por fim, devido ao fato de haver tido na região diversos períodos chuvosos, observou-se um crescente número de *downloads* do aplicativo em dispositivos móveis da população.

A este aplicativo tem acesso um total de 332.433 habitantes, atendendo a 4 municípios, com o monitoramento de 177 Km do Rio Itajaí Mirim. Os municípios atendidos são:

- Vidal Ramos: 6 293 habitantes;
- Botuverá: 4 864 habitantes;
- Brusque: 119 719 habitantes;
- Itajaí: 201 557 habitantes.

Justificativa para incluir a acessibilidade no Alerta Brusque

Estudos da área de usabilidade nos comprovam que, em frações de segundos, se retêm ou expelle usuários dos dispositivos móveis. Considerando que o Alerta Brusque, é um aplicativo de referência e em crescente uso pela população do rio Itajaí Mirim e um importante recurso de orientação junto à Defesa Civil, torna-se evidente a necessidade de melhorar seu acesso e navegação.

Na versão atual disponível, as diretrizes de acessibilidade (Lei 5.296) não são atendidas, o usuário perde tempo ao tentar localizar manualmente as informações. A lei nos diz que todos os dispositivos devem oferecer recursos que exijam menor esforço físico e mental dos usuários, minimizando sua desorientação ou sobrecarga cognitiva.

Em situações de emergência, é fundamental a fácil interação dos usuários com o aplicativo para localizar as informações. Portanto, é importante pesquisar para implantar no Alerta Brusque respostas rápidas por meio de recursos tecnológicos e técnicas de Inteligência Artificial ao invés de utilizar como principal forma de entrada somente a *touch-screen* e teclado, perdendo tempo para encontrar o que precisa. Devido à evolução das tecnologias e a computação ubíqua, gesto e voz possibilitam criar aplicações extremamente ricas, especialmente para pessoas com pouca experiência ou com necessidades especiais.

Tratando-se de sistemas de emergência, como é o caso do Alerta Brusque, é fundamental melhor usabilidade para permitir rápido e fácil acesso aos níveis de rio e chuva em situação de preocupação extrema.

Como alternativa mais natural para o usuário interagir facilmente com o Alerta Brusque e obter imediatismo das informações, a técnica por comando de voz e resposta em áudio é proposta, ao invés de o usuário ter que ler no dispositivo móvel e selecionar manualmente suas informações.

Projetar aplicativos que executam tarefas com mãos e olhos livres mostra-se extremamente vantajoso, independente do tipo de usuário, ele será capaz de executar o sistema com apenas alguns segundos de orientação das atividades através de comandos sonoros e confirmação das solicitações verbalmente. Estudos feitos no segmento de logística mostraram-se muito eficientes ao reduzir em até 85% os erros em operações, além de garantir 100% de rastreabilidade sobre as mercadorias que circulam no depósito (BERNARDES, s.d.).

Em situações de risco causadas por desastres ambientais, percebe-se cada vez mais a importância de informações precisas divulgadas de forma a passar confiança. É perceptível a

falha na gestão da informação durante desastres naturais no município de Brusque e no Vale do Itajaí em geral. Muitos órgãos, como Defesa Civil, não possuem estrutura suficiente para atender demandas de acesso ao seu *site*, ocorrendo quedas constantes nos momentos em que a população mais precisa fazer acompanhamento do nível de rio e precipitação de chuva.

Fundamentação teórica

Os avanços nas técnicas de reconhecimento de voz indicam viabilidade da tecnologia em diversas aplicações, sobretudo em dispositivos móveis. Segundo Hearst (2011), alguns fatores contribuem para a demanda de interfaces com reconhecimento de voz nesses dispositivos: (1) por se tratar de um caminho natural ao uso da fala e (2) as interfaces de toque dificultam a escrita de textos longos.

Existem tecnologias de reconhecimento de voz maduras que não necessitam de treinamento, como *Apple Siri*, *Microsoft Cortana*, *GoogleNow*. Três fatores contribuíram para essa maturidade: o aumento da capacidade computacional, a quantidade de dados disponíveis para treinamento e finalmente a evolução dos algoritmos de aprendizado de máquina (PICHENY, 2015). É possível identificar padrões complexos de voz com baixo tempo de resposta, utilizando dispositivos portáteis e navegadores *web*. Isso sugere que o uso da voz deve aumentar significativamente à medida que o tempo de resposta e a precisão melhoram a ponto de atender às necessidades reais dos usuários (HEARST, 2011).

Aplicações legadas no modelo cliente-servidor estão evoluindo para soluções *web*, onde a usabilidade é considerada um fator de sucesso do produto (HAINE 2012). Dessa forma, um modelo de interação mais natural utilizando a voz pode aumentar a competitividade no mercado.

Portanto, será pesquisado sobre a *engine* do Google através da *Web Speech API* para reconhecimento de voz, visando construir uma solução que atenda à demanda dos usuários conforme levantado por Schnelle et al. (2005). A escolha se baseou no suporte ao português do Brasil e na portabilidade. Atualmente, a API é limitada ao navegador Google Chrome, mas é uma implementação da especificação da *W3C* (GLEN SHIRES, 2012), podendo ser adotada pelos demais navegadores.

A conversão de texto em voz é outra ótima ferramenta para tornar mais prática a relação do usuário com os dispositivos móveis. Afinal, pode avisar, em voz alta, as solicitações feitas pelo usuário e ler mensagens. Leitores de telas podem ser utilizados para responder as opções do Alerta Brusque obtendo retorno em áudio sobre o próprio texto apresentado.

Apesar do investimento muito substancial na pesquisa da tecnologia de reconhecimento de fala nos últimos 40 anos, o sintetizador de fala e as tecnologias do reconhecimento de fala têm ainda limitações significativas (MARANGONI; PRECIPITO, 2006). O mais importante, o sintetizador de fala não se encontra sempre com as expectativas elevadas dos usuários familiares com uma comunicação de fala humano a humano-natural. Compreender as limitações é importante para o uso eficaz da entrada e da saída da fala em uma relação de usuário.

Por fim, segundo Melo e Pupo (2010), entre os dispositivos de voz sintetizada, estão os programas que convertem texto em fala (ex.: *DeltaTalk*) e os leitores de tela com síntese de voz (ex.: *Jaws for Windows*, *NVDA*, *Orca*, *Virtual Vision*). Com o *DeltaTalk*, desenvolvido para sistema Windows, o usuário seleciona um texto e aciona a tecla virtual <F9> para que este seja "falado". Já os leitores de tela, além de converterem texto em fala, captam as informações textuais exibidas na tela do computador e as apresentam utilizando voz sintetizada.

Trabalhos Relacionados

A literatura sobre técnicas de reconhecimento de voz é abrangente e teve evoluções significativas desde o início do século 21 (PICHENY, 2015). No entanto, trabalhos sobre a utilização e desempenho dos mecanismos mais modernos de reconhecimento, como a API do Google para português do Brasil são ainda pouco explorados. A seguir, os principais trabalhos utilizando a abordagem de reconhecimento de comandos com a *Web Speech* API são discutidos.

Em Skraba et al. (2014), é demonstrada a utilização do reconhecimento de voz no controle de uma cadeira de rodas. A abordagem considera o *Web Speech* API do Google e possui uma limitação análoga ao trabalho atual com a necessidade de conexão com a internet. Porém, essa limitação é mais severa para a cadeira de rodas devido à mobilidade. Uma diferença em relação à presente proposta está na complexidade dos comandos, uma vez que os comandos de controle da cadeira de rodas são simples e curtos. Finalmente, há uma diferença significativa no objetivo, uma vez que buscamos uma validação com usuários finais enquanto tinham um objetivo em propor uma solução de software e hardware (SKRABA et al., 2014).

O trabalho realizado em Kimura et al. (2015) é uma aplicação *web* para realizar testes de pronúncia para estudantes não nativos da língua inglesa. Os pesquisadores utilizam o *Web Speech* API do Google para salvar a transcrição e comparar com a resposta do material de ensino. Um desafio comum ao trabalho mencionado é a necessidade de comparar o resultado do reconhecimento de voz com um resultado esperado. A estratégia utilizada Kimura et al. (2015) foi a comparação de *arrays* enquanto o presente trabalho utiliza expressões regulares. Finalmente, o trabalho não aborda o reconhecimento de voz em português do Brasil, o que caracteriza um desafio, uma vez que os principais *engines* do mercado são desenvolvidos para o inglês.

Um sistema autônomo para suporte a idosos é proposto por Valencia et al. (2014). O sistema realiza o reconhecimento de voz com dois microfones e uma técnica de pré-processamento *fuzzy* para melhorar a captação. A principal diferença em relação ao trabalho atual é o modelo de entrada de áudio que possui captação de dois microfones com tratamento antes do envio ao *Web Speech* API. No presente trabalho, utilizamos uma abordagem de captação simples para viabilizar a utilização em qualquer ambiente com um navegador e com microfone.

A revisão da literatura indica que a utilização do *Web Speech* API é considerada pelos pesquisadores citados como uma alternativa viável a interfaces naturais. No entanto, os problemas encontrados e suas soluções quanto ao mapeamento de reconhecimento de voz para comandos é pouco detalhado. São esses tópicos que exploramos com este trabalho.

No *Windows 10*, é possível ouvir o texto em voz alta através do agente Narrador. O sintetizador de voz lê o texto na tela do computador em voz e descreve eventos, como notificações ou compromissos do calendário, para que você possa usar seu computador sem um vídeo. Para iniciar ou fechar o Narrador, pressione a tecla do logotipo do *Windows* + Enter. O Narrador está disponível em inglês (Estados Unidos, Reino Unido e Índia), francês, italiano, alemão, japonês, coreano, mandarim (chinês simplificado e chinês tradicional), cantonês (chinês tradicional), espanhol (Espanha e México), polonês, russo e português (Brasil) (*WINDOWS 10*, s.d.).

Os recursos sintetizadores de voz do *Windows* utilizam a biblioteca *System.Speech*, junto com a voz correspondente ao idioma, para converter o texto em áudio.

Segundo Álvares (s.d.), a biblioteca *System.Speech*, do *.net framework*, possui 3 poderosas *namespaces*: (1) *System.Speech.Synthesis: Namespace* responsável por conter classes para manipular o sintetizador de voz, capaz de transformar texto em áudio. (2) *System.Speech.AudioFormat: Namespace* responsável por conter classes para gerar o arquivo de áudio e definir a qualidade dele. (3) *System.Speech.Recognition: Namespace* contém classes e *namespaces* responsáveis por

fazer o efeito inverso do proposto neste artigo; ou seja, reconhecer a voz e transformar em texto.

Outra ferramenta é a *ReadSpeaker TTS Production* API que oferece um método automatizado para gerar arquivos áudio, dando acesso a vozes de alta qualidade em muitos idiomas e permite que a sua aplicação envie texto e receba os correspondentes arquivos de áudio gerados pelo serviço de texto-voz (READSPEAKER, 2016). E o *JAVA Speech* API conversa com o *IBM ViaVoice*.

Por fim, o *Java Speech* API define um padrão para a utilização da fala para interação com o computador. Duas tecnologias de fala são suportadas pelo *Java Speech* API. Uma delas é o reconhecimento de fala (reconhecimento da fala) e a outra é o sintetizador de fala (síntese da fala). O reconhecimento de fala fornece aos computadores a habilidade de “escutar” a língua falada e de determinar o que foi dito, ou seja, processa a entrada de áudio que contém a fala convertendo para texto. O sintetizador de fala, por sua vez, fornece o processo reverso de produzir a fala sintética do texto gerado por uma aplicação, por um *applet* ou por um usuário. É chamada frequentemente como a tecnologia *text-to-speech* (texto para fala).

Considerações finais

Buscou-se, com o presente artigo, extrair o máximo de conhecimento sobre os métodos e técnicas de programação das APIs (*Application Programming Interface*) para comando por voz e orientação em áudio oferecendo melhor usabilidade e facilidade de uso no Alerta Brusque, uma nova forma de manipulação e acessibilidade às funcionalidades sobre previsões e prevenções dos níveis do rio e chuva do rio Itajaí Mirim.

As contribuições científicas e tecnológicas desta proposta são diversas, pois a utilização da Inteligência Artificial está presente em todas as áreas. Seus benefícios são incalculáveis quando aplicados no auxílio da população como meio de atenção e orientações seguras diante de momentos de enchentes e deslizamentos. A região que o rio Itajaí Mirim percorre não possuía automatização em seus processos hidrológicos em auxílio à Defesa Civil de Brusque em Santa Catarina.

O sintetizador de voz do Alerta Brusque irá interpretar a fala do usuário e buscar resposta à solicitação feita ao aplicativo e converterá a informação textual em fala (áudio) para o usuário. Portanto, a partir desta inovação, o grau de aceitação do aplicativo deverá aumentar significativamente, haja vista que não existem na área de desastres naturais recursos inteligentes e interativos com este nível de usabilidade.

Referências

ÁLVARES, R. **Transformando texto escrito em texto narrado com .NET e a biblioteca System.Speech**. Disponível em: <<http://www.devmedia.com.br/transformando-texto-escrito-em-texto-narrado-com-net-e-a-biblioteca-system-speech/24240>>. Acesso em: 21 mar. 2016.

BERNARDES, W. K. C. **Tecnologia de voz ganha espaço no setor de logística**. Disponível em: <<http://www.cgimoveis.com.br/logistica/tecnologia-de-voz-ganha-espaco-no-setor-de-logistica>>. Acesso em: 24 mar. 2016.

BRASIL. Lei Nº 5.296, de 2 de Dezembro de 2004. Regulamenta as Leis nºs 10.048, de 8 de novembro de 2000, que dá prioridade de atendimento às pessoas que especifica, e 10.098, de 19 de dezembro de 2000, que estabelece normas gerais e critérios básicos para a promoção da acessibilidade das pessoas portadoras de deficiência ou com mobilidade reduzida, e dá

outras providências. 2004. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/decreto/d5296.htm>. Acesso em: 9 abr. 2016.

GLEN SHIRES, H. W. **Web speech API specification**. 2012. Disponível em: <<https://dvcs.w3.org/hg/speech-api/raw-file/tip/speechapi.html>>. Acesso em: 19 mar. 2016.

HAINED, P. **The ux revolution at successfactors**. 2012. Disponível em: <<http://www.successfactors.com/static/docs/successconnect/london/successfactors-philip-haine.pdf>>. Acesso em: 19 mar. 2016.

HEARST, M. A. 'Natural' search user interfaces. 2011. *Commun. ACM*, v. 54 (11): p. 60-67, 2011.

KIMURA, H. et al. The effects of listening agent in speech-based on-line test system. **Global Engineering Education Conference (EDUCON)**, IEEE, p. 366-370, 2015.

MARANGONI, J. B.; PRECIPITO, W. B. Reconhecimento e sintetização de voz usando java speech. **Revista científica eletrônica de sistemas de informação**. Ano II, Número 4, 2006. Disponível em: <http://faef.revista.inf.br/imagens_arquivos/arquivos_destaque/bjMnA-2Zwc9685z8_2013-5-27-15-40-25.pdf>. Acesso em: 20 mar. 2016.

MELO, A. M.; PUPO, D. T. **A Educação Especial na Perspectiva da Inclusão Escolar: livro acessível e informática acessível**. Brasília: Ministério da Educação, Secretaria de Educação Especial. Universidade Federal do Ceará, 2010. Disponível em: <http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=7119-fasciculo-8-pdf&category_slug=novembro-2010-pdf&Itemid=30192>. Acesso em: 20 mar. 2016.

PICHENY, M. **Ibm watson now brings cognitive speech capabilities to developers**. 2015. Disponível em: <<https://goo.gl/Pmzpb0>>. Acesso em: 19 mar. 2016.

READSPEAKER. The voice of the web. Produção Áudio. Disponível em: <<http://www.readspeaker.com/pt-pt/producao-audio/>>. Acesso em: 20 mar. 2016.

SCHNELLE, D. et al. Audio navigation patterns. **EuroPLoP**, p. 237-260, 2005.

SKRABA, A. et al. Prototype of speech controlled cloud based wheelchair platform for disabled persons. In: **Embedded Computing (MECO)**. 3rd Mediterranean Conference on, p. 162-165, 2014.

VALENCIA, R. et al. Sa3m: An interactive robot to provide support for the elderly. **Power, Electronics and Computing (ROPEC)**, IEEE International Autumn Meeting on, p. 1-6, 2014

WINDOWS 10. **Ouçã o texto lido em voz alta com o Narrador**. Disponível em: <<http://windows.microsoft.com/pt-br/windows-10/getstarted-hear-text-read-aloud>>. Acesso em: 20 mar. 2016.

WAZLAWICK, Raul Sidney. **Metodologia de Pesquisa para Ciência da Computação**. Rio de Janeiro: Elsevier, 2008.

Artigo recebido em 15/06/16. Aceito em 18/08/16.