



Proposta de Serviço Piloto

Grupo de Trabalho – Segunda Fase

GTAaaS – Acessibilidade como um Serviço (Accessibility as a Service)

Prof. Dr. Guido Lemos de Souza Filho (UFPB)
05 de Setembro de 2012

1. Concepção do serviço

1.1. Resumo

Pessoas com deficiência enfrentam sérias dificuldades para acessar informações. Além das instituições educacionais não estarem preparadas para lidar com esse público, as

tecnologias de informação e comunicação (TIC) quando são desenvolvidas raramente levam em conta os seus requisitos. Para reduzir esses problemas, neste projeto é explorado o conceito de “Acessibilidade como um Serviço”, através do projeto e desenvolvimento de um serviço para viabilizar o acesso de usuários surdos a conteúdos digitais multimídia. Na primeira fase do projeto, um protótipo do serviço proposto foi desenvolvido e, em seguida, validado através de um conjunto de testes com usuários surdos. Neste protótipo, um vídeo em Língua Brasileira de Sinais (LIBRAS) é gerado automaticamente, a partir da tradução da legenda do conteúdo multimídia original, e é embarcado nesse conteúdo utilizando um processo de mixagem. Na segunda fase, a proposta do projeto é integrar o serviço piloto com o serviço de Vídeo sob Demanda (VoD) da RNP, permitindo que os conteúdos disponíveis nos repositórios desse serviço possam ser disponibilizados em formato acessível aos clientes da RNP. Outra alternativa seria integrar o serviço piloto em outros serviços da RNP, como o “Videoaula@RNP”, o serviço de “Transmissão de vídeo ao vivo” ou o “serviço de Transmissão de sinal de TV”, permitindo assim que outras demandas (por conteúdos digitais acessíveis) dos clientes da RNP também possam ser atendidas. Por fim, novas funcionalidades e estratégias de computação humana serão incluídas no serviço para permitir a revisão das traduções produzidas, bem como a geração das legendas dos conteúdos e o desenvolvimento eficiente do Dicionário de LIBRAS.

1.2. Abstract

People with disabilities have serious difficulties to access information. Educational institutions are not well prepared to address needs of these users. In addition, the information and communication technologies (ICT) are rarely developed look on their specific requirements. To minimize this problem, in this project we explore the concept of “Accessibility as a Service” by providing and developing a service that makes it possible for deaf users to access digital multimedia contents. In the first phase of the project, a prototype of the proposed service was developed and validated through a set of tests with Brazilian deaf users. In this prototype, a Brazilian sign language (LIBRAS) video is automatically generated from the translation of subtitles of the original multimedia content and then mixed to it. In the second phase, the proposal is to integrate the pilot service with the RNP Video on Demand (VoD) service, making possible it contents be also available in an accessible format. Another alternative is to integrate the pilot service in other RNP services, such as the “Videoaula@RNP” service, the “Live video transmisson service or the “TV signal transmission” service, which allow that other RNP customers demands (for digital content accessible) can be met. Finally, new features and human computation strategies will be included in the prototype to enable the revision of the translations produced by the service, as well as the generation of the subtitles and the development of the LIBRAS Dictionary.

1.3. Descrição do serviço proposto

Os surdos representam uma parcela significativa da população brasileira e mundial. Segundo o censo demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE) de 2010, existem, no Brasil, cerca de 9,5 milhões de pessoas com algum nível de deficiência auditiva, o que representa cerca de 5,1% da população brasileira [IBGE 2010]. Em termos mundiais, a estimativa da Organização Mundial de Saúde (OMS) é de que aproximadamente 275 milhões de pessoas apresentem algum nível de deficiência auditiva [WHO 2012]. Esses indivíduos, no entanto, enfrentam uma série de dificuldades para acessar informações. Eles se comunicam naturalmente através de línguas gestuais, denominadas línguas de sinais, e, em virtude disso, as línguas orais utilizadas cotidianamente pela maioria das pessoas e em praticamente todos os meios de comunicação, quando conhecidas, representam apenas “uma segunda língua” [Campos & Girafa 2000].

Em consequência disso, eles enfrentam grandes dificuldades para compreender e se comunicar através de textos em línguas orais. Uma vez que essas línguas possuem uma grafia baseada em sons, muitos deles passam vários anos na escola e não conseguem aprender a ler e escrever na língua oral de seu país [Stumpf 2000]. No Brasil, por exemplo, segundo o censo demográfico do IBGE de 2010 [IBGE 2010], cerca de 75% dos deficientes não concluem o ensino médio. Um outro exemplo que pode ser citado é que um estudo realizado por Wauters (2005) com crianças e adolescentes surdos holandeses de 7 a 20 anos de idade, mostra que apenas 25% deles possuem uma capacidade de leitura igual ou superior ao de uma criança sem deficiência de 9 anos.

As Tecnologias da Informação e Comunicação (TIC), quando desenvolvidas, raramente levam em conta os requisitos e necessidades deste público [Haddon & Paul 2001]. O suporte para línguas de sinais, na TV, por exemplo, é, em geral, limitado aos dispositivos manuais, onde uma janela com um intérprete de língua de sinais é apresentada junto com o vídeo original do programa (*wipe*). Essa solução, além de possuir altos custos operacionais para geração e produção (câmeras, estúdio, equipe, etc.) dos conteúdos, necessita de intérpretes humanos em tempo integral, o que acaba restringindo seu uso a uma pequena parcela da programação. Essas dificuldades resultam em uma grande barreira para a comunicação dos surdos com outras pessoas, o acesso a informações, a aquisição de conhecimentos, dentre outros.

Nas instituições educacionais de ensino, também existem barreiras importantes para o acesso desses usuários à informação. Embora essas instituições sejam obrigadas por lei a dar suporte e assistência especial aos estudantes surdos, não existem intérpretes de LIBRAS (Língua Brasileira de Sinais) disponíveis para todos os estudantes/disciplinas, o que acaba reduzindo o acesso aos conteúdos ministrados nas aulas. Essas dificuldades se tornam ainda mais evidentes, considerando que vêm crescendo, nos últimos anos, o número de estudantes com surdez que estão ingressando no ensino técnico e superior [IBGE 2010].

Para minimizar esses problemas, uma solução alternativa seria viabilizar a geração automática de vídeos em língua de sinais, a partir da tradução automática de legendas de conteúdos digitais multimídia (como, por exemplo, vídeo-aulas, conteúdos de palestras, vídeos da Web e Cinema Digital, sinal de TV, dentre outros). Dentro deste contexto, neste projeto é explorado o conceito de “Acessibilidade como um Serviço”, que tem como objetivo projetar e desenvolver um serviço para geração automática de conteúdos acessíveis para surdos. A proposta do serviço é que usuários submetam vídeos e o serviço piloto gere automaticamente uma janela em Língua Brasileira de Sinais (LIBRAS) para estes vídeos a partir da sua legenda, tornando-os acessíveis para os surdos brasileiros. Com isso, é possível reduzir significativamente os custos de geração e produção das janelas de LIBRAS, viabilizando, dessa forma, o acesso a estes conteúdos para os usuários surdos nos mais diversos contextos (vídeo aulas, palestras, TV, Web, Cinema Digital, dentre outros).

Em consequência disso, um dos principais benefícios do serviço proposto é a possibilidade de disponibilizar conteúdos digitais acessíveis aos clientes RNP. Uma vez que a RNP possui um conjunto de serviços de “Disponibilização de Conteúdos Digitais” (Vídeo sob demanda, Vídeoaula@RNP, Transmissão de Sinal de TV e Transmissões de Vídeo ao Vivo), a integração desses serviços com o serviço proposto permitiria que esses conteúdos também fossem disponibilizados em formato acessível para surdos, viabilizando a inclusão desses usuários nos serviços providos pela RNP.

Em virtude disso, nessa segunda fase do projeto, a proposta inicial é integrar o protótipo desenvolvido na primeira fase com o serviço de Vídeo sob Demanda (VoD) da RNP, permitindo que os conteúdos disponíveis nos repositórios desse serviço possam ser disponibilizados em formato acessível aos clientes da RNP. Opcionalmente, no entanto, também seria possível integrar o serviço piloto aos outros serviços de Disponibilização de Conteúdos Digitais da RNP, como o “Vídeoaula@RNP”, o serviço de “Transmissão de vídeo ao Vivo” ou o serviço de “Transmissão de Sinal de TV”. Adicionalmente também serão incluídas novas funcionalidades e estratégias de computação humana no protótipo piloto desenvolvido na primeira fase, com o objetivo de melhorar a qualidade das traduções produzidas. Dentre as funcionalidades, pode-se destacar: a possibilidade de revisar as traduções produzidas pelo serviço, a exploração de estratégias de computação humana para a legendagem dos vídeos e o desenvolvimento eficiente do Dicionário de LIBRAS. Por fim, o Dicionário de LIBRAS produzido pela solução poderá ser armazenado no FEB da RNP para disponibilização em outras ferramentas e serviços da RNP.

1.4. Identificação do público alvo

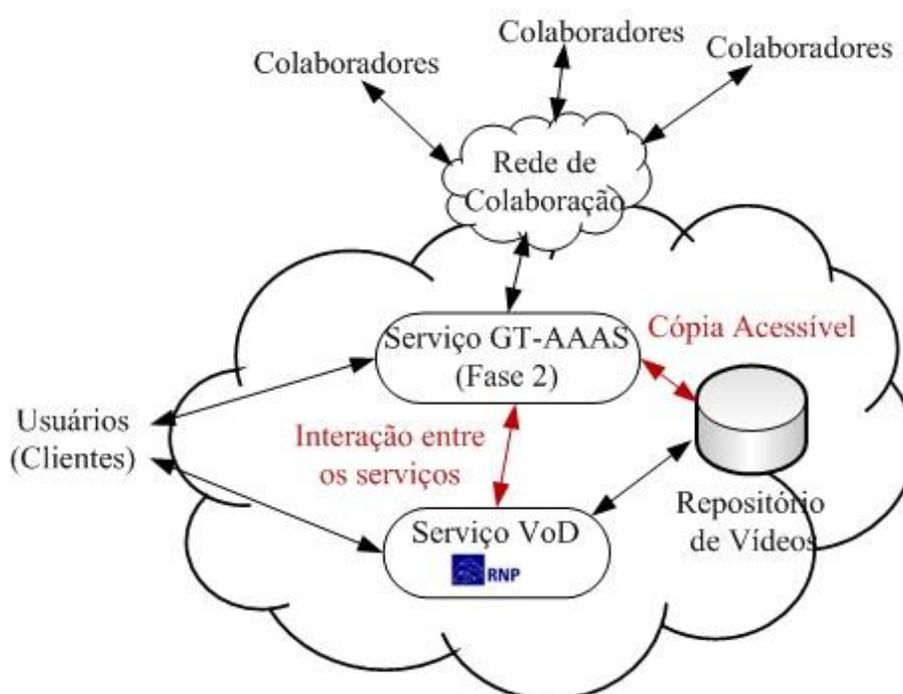
O público alvo do serviço proposto são todos os clientes da RNP que tenham interesse em produzir conteúdos digitais acessíveis (Instituições de Ensino e Pesquisa, Instituto Nacional de Educação de Surdos – INES –, TVs Universitárias, dentre outros), além dos próprios usuários surdos das instituições clientes da RNP.

2. Definição do serviço piloto

2.1. Arquitetura do serviço piloto

Conforme mencionado anteriormente, nesta segunda fase do projeto, a proposta inicial é que o protótipo do GT-AAAS (desenvolvido na primeira fase do projeto) seja refinado e integrado ao serviço de Vídeo sob Demanda (VoD) da RNP. Com isso, é possível que os clientes da RNP possam produzir cópias acessíveis dos seus conteúdos (vídeos) armazenados nos repositórios do serviço de VoD da RNP. Na Figura 1, é ilustrada uma visão esquemática do serviço piloto resultado dessa integração.

Figura 1. Visão esquemática da integração com o serviço de VoD.

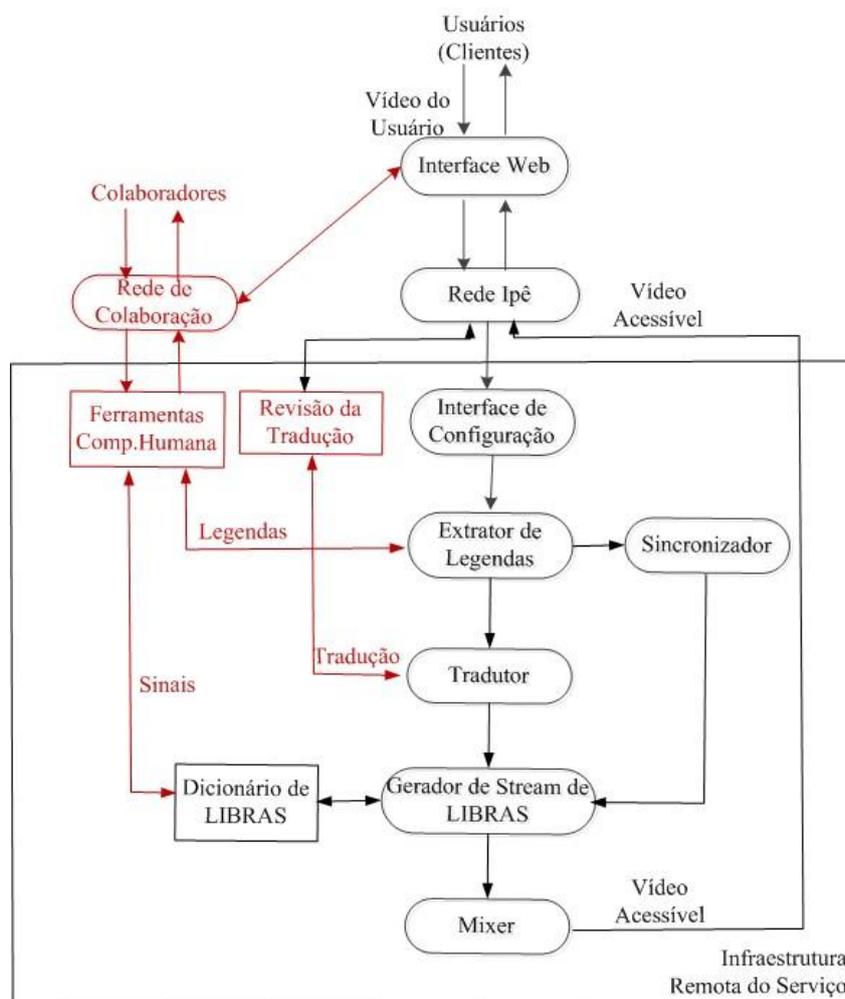


De acordo com a Figura 1, o serviço piloto funcionará da seguinte forma. Inicialmente, um usuário (cliente) submete um vídeo para o serviço piloto. Essa submissão pode ser feita em dois cenários: (1) diretamente, utilizando a interface do serviço do GT-AAAS ou (2) indiretamente, utilizando o serviço de VoD. No primeiro cenário, o usuário submete o vídeo diretamente utilizando a interface do serviço do GT-AAAS, e, nesse caso, o processo de geração do conteúdo acessível é similar ao processo utilizado no protótipo da primeira fase: um vídeo em Língua Brasileira de Sinais (LIBRAS) é gerado automaticamente, a partir da tradução da legenda do conteúdo multimídia do usuário, e é embarcado nesse conteúdo utilizando um processo de mixagem. No segundo cenário, o serviço de VoD funcionaria como um intermediário para a geração da cópia acessível. Nesse caso, um parâmetro de configuração seria incluído na interface de configuração do VoD para permitir a geração de vídeos acessíveis. Caso esse parâmetro seja configurado, o VoD invoca (interage com) o serviço do GT-AAAS para que esse produza

uma **cópia acessível** do conteúdo selecionado pelo usuário. A **cópia acessível** gerada pelo serviço do GT-AAAS é então devolvida ao VoD que a armazena no repositório.

Adicionalmente, estratégias de computação humana serão incluídas na implementação do serviço piloto para auxiliar em dois elementos chaves no processo de geração dos vídeos acessíveis: a geração das legendas dos vídeos (elemento base para o processo de tradução) e a geração do Dicionário de LIBRAS (elemento base para geração do vídeo de LIBRAS). Somado a isto, uma estratégia de revisão da tradução também será incluída na arquitetura do serviço do GT-AAAS para permitir que os usuários revisem as traduções produzidas pelo serviço. A arquitetura do serviço do GT-AAAS com a inclusão destas novas funcionalidades (destacadas em vermelho) é apresentada na Figura 2.

Figura 2. Arquitetura do serviço do GT-AAAS.



De acordo com a Figura 2, inicialmente, o usuário acessa o serviço através de uma Interface Web. Nesta interface, o usuário submete o vídeo a ser traduzido e configura alguns parâmetros relacionados a janela de LIBRAS gerada no processo de tradução como, por exemplo, o tamanho, a posição, a cor do fundo da janela de LIBRAS, dentre outros. O vídeo, juntamente com os parâmetros de configuração da janela de LIBRAS, são então encaminhados para a infraestrutura remota do serviço. Na infra-estrutura

remota do serviço, um componente **Interface de Configuração** configura e invoca os componentes do serviço. Um componente **Extrator de Legendas** é então aplicado sobre o fluxo de legenda do vídeo para convertê-lo em um conjunto de sentenças em língua portuguesa. Em seguida, essas sentenças são automaticamente traduzidas (**Tradutor**) para uma seqüência de *glosas* (representação textual de LIBRAS), que são convertidas pelo componente **Gerador do Stream de LIBRAS** em um vídeo de LIBRAS com o auxílio do **Dicionário de LIBRAS**. Os quadros do vídeo de LIBRAS são então mixados com os quadros do vídeo original do usuário (**Mixer**), tornando assim o vídeo acessível para usuários surdos e devolvendo-o ao usuário.

A sincronização entre a entrada (legenda ou áudio em língua portuguesa) e a saída (vídeo de LIBRAS) é realizada utilizando o modelo de sincronização baseado no eixo do tempo [Blakowski & Steinmetz 1996]. Esse modelo define pontos de sincronização que são inseridos no fluxo usando marcações de tempo (*timestamps*) baseadas em um relógio global. Nesse caso, o relógio global é o relógio do arquivo de vídeo. Dessa forma, é possível inserir marcações de tempo no vídeo de LIBRAS de forma sincronizada com as marcações das legendas.

Em paralelo com o processo de geração do vídeo acessível, um conjunto de ferramentas serão incluídas no protótipo para melhorar o processo de geração. Esses componentes destacados em vermelho, incluem uma ferramenta de **Revisão da Tradução**, para que os usuários possam revisar as traduções geradas pelo serviço, e **Ferramentas de Computação Humana**, para geração das legendas dos vídeos e dos sinais do Dicionário de LIBRAS por usuários colaboradores.

Adicionalmente, o componente **Interface de Configuração** será ajustado/adaptado para permitir a interação (comunicação) com outros serviços da RNP, como, por exemplo, com o serviço de VoD. Para isso, um protocolo de comunicação será definido para o serviço do GT-AAAS. Com esse protocolo será possível definir uma estratégia padrão de integração do serviço do GT-AAAS com qualquer outro serviço da RNP que envolva “Disponibilização de Conteúdos Digitais”.

2.2. Instituições participantes

As instituições envolvidas no desenvolvimento desse serviço piloto serão o LAViDUFPB, a Fundação CPqD (respectivamente, os executores e co-executores do projeto GT-AAAS na Fase 1) e o LARC-USP que cederá uma cópia do código do serviço de VoD da RNP. Adicionalmente, o piloto será validado para um conjunto de vídeos dentro do escopo das instituições que são clientes do serviço de VoD da RNP. Essas instituições e seus contatos são apresentadas abaixo:

- **LAVID – Núcleo de Pesquisa e Extensão Laboratório de Aplicações em Vídeo Digital**
Contatos: Guido Lemos S. Filho e Tiago Maritan U. de Araújo
E-mail: guido@lavid.ufpb.br e tiagomaritan@lavid.ufpb.br
- **CPqD – Centro de Pesquisa e Desenvolvimento em Telecomunicações**
Contato: Lara Schibelsky Godoy Piccolo

E-mail: lpiccolo@cpqd.com.br

- **LARC – Laboratório de Arquitetura e Redes de Computadores**
Contato: Regina Silveira
Contato: regina@larc.usp.br

2.3. Refinamento do protótipo

Conforme mencionado na Seção 2.1, o protótipo do GT-AAAS será refinado com a inclusão de novas funcionalidades como, por exemplo, uma ferramenta de **Revisão da Tradução** para que os usuários possam revisar as traduções geradas pelo serviço, e a inclusão de **Ferramentas de Computação Humana** para legendagem dos vídeos do repositório e para geração colaborativa do Dicionário de LIBRAS. Além disso, o componente **Interface de Configuração** será ajustado para permitir a interação com outros serviços da RNP, usando um protocolo padrão de comunicação que será desenvolvido nessa segunda fase.

Adicionalmente, algumas funcionalidades de suporte a operação também serão incluídas para viabilizar a administração da ferramenta, cadastro e autenticação de usuários, dentre outros. Essas ferramentas serão descritas na Seção 2.4.

2.4. Ferramentas de suporte à operação

No protótipo desenvolvido na primeira fase do projeto, além de submeter os vídeos e fazer download das cópias acessíveis produzidas pelo protótipo, os usuários configuravam basicamente os parâmetros de tamanho e posição da janela de LIBRAS e o nível de transparência do *background* dessa janela (opaco ou transparente). Para viabilizar a operação e gerência do serviço do GT-AAAS, nessa segunda fase do projeto, algumas funcionalidades de suporte a operação e gerência do serviço também serão incluídas no protótipo. Essas funcionalidades envolvem, principalmente, o cadastro e autenticação dos usuários e uma interface de administração do serviço e são melhor descritas a seguir:

- **Cadastro e Autenticação dos Usuários:** Envolve o cadastro dos usuários no banco de dados (BD) do serviço e a sua autenticação para uso da ferramenta.
- **Interface de administração do serviço:** Permitirá monitorar os usuários e grupos cadastrados, monitorar as estatísticas de uso do serviço, o número de sinais e legendas produzidos nas ferramentas de computação humana, monitorar os servidores e processos, dentre outros.

3. Cronograma

O cronograma de atividades proposto para a segunda fase do GT-AAAS é apresentado na Tabela 1.

Tabela 1. Cronograma de Atividades da Segunda Fase do GT-AAAS

Atividades	1º Trim.	2º Trim.	3º Trim.	4º Trim.
1. Desenvolvimento do Protocolo de Comunicação do Serviço do GT-AAAS				
2. Ajustes no componente Interface de Comunicação				
3. Desenvolvimento e Integração da Ferramenta de Revisão da Tradução				
4. Desenvolvimento e Integração das Ferramentas de Computação Humana (Geração de Legendas e Sinais)				
5. Desenvolvimento das Ferramentas de Suporte a Operação do Serviço				
6. Integração com o serviço de VoD da RNP				
7. Testes do serviço piloto (GT-AAAS + VoD)				
8. Redação de artigos científicos com os principais resultados do projeto.				

4. Recursos financeiros

4.1. Equipamentos e softwares

Nome	Descrição	Qde.	Valor Unitário	Valor Total
Desktops	Dell Vostro 260 Slim com Processador Intel i5-2400 (3.1GHz); 6GB de RAM DDR3 (1333MHz); 1GB SATA (7200 RPM); Monitor Dell 18,5"; Mouse e teclado. (Para desenvolvimento)	4	R\$ 2.500,00	R\$ 10.000,00
MiniPC	MiniPC AOpen DE67-HA com Processador Intel i7 – 2,7GHz – 5.0GT/s – 4MB ; 8GB de RAM DDR3 – 1333MHz; 1TB 5400RPM SATA 2; 802.11a/b/g/n (Utilizado para viagens e demonstrações do protótipo)	1	R\$ 4.000,00	R\$ 4.000,00
Memória	Memória de 8GB (1x8GB, 2Rx4, 1.35V) PC3L-10600 CL9 ECC DDR3 1333MHz LP RDIMM para Servidor IBM x3650 (Composição do Servidor do piloto)	2	R\$ 750,00	R\$ 1.500,00
Disco de Armazenamento	HD 1TB 2.5in SFF Hot Swap 7.200 RPM 6Gbps NL SATA para Servidor IBM x3650 (Composição do servidor do piloto)	2	R\$ 1.800,00	R\$ 3.600,00
Câmera Filmadora Sony	Filmadora Full HD HDR-CX190 (Gravação de vídeos da interface do sistema e testes do sistema)	1	R\$ 1.600,00	R\$ 1.600,00
Câmera Fotográfica Sony	Câmera Digital 16.1MP - DSC-W630 (Testes e avaliação do sistema)	1	R\$ 500,00	R\$ 500,00
Tablet	Tablet Motorola Xoom Android 3.0 (Testes com usuários surdos e demonstrações em conjunto com o GT de Visualização Avançada)	2	R\$ 2000,00	R\$ 4.000,00

Total				R\$ 25.200,00
--------------	--	--	--	---------------

4.2. Viagens Nacionais

Origem	Destino	Objetivo	Nº Passagens	Nº Diárias
JPA	VCP	Reunião para refinamento do protótipo	01	02
VCP	JPA	Reunião de acompanhamento do projeto	02	04
JPA	CGH	Reunião para integração com VoD	03	06

5. Referências

BLAKOWISKI, G.; STEINMETZ, R.. A Media Synchronization Survey: Reference Model, Specification and Case Studies. IEEE Journal on Selected Areas in Communication, 14, 1996. 5-35.

CAMPOS, M. B.; GIRAFA, L. M. M. **SIGNSIM**: uma ferramenta para auxílio à aprendizagem da língua brasileira de sinais. V Congresso Ibero-Americano de Informática na Educação – RIBIE. Viñadelmar, Chile: [s.n.]. 2000. p. 1-13.

HADDON, L.; G. PAUL. Technology and the Market: Demand, Users and Innovation, ASEAT Conference Proceedings Series, Edward Elgar Publishing, Cheltenham, Reino Unido, capítulo Design in the ICT industry: the role of users. 2001, pp. 201–215.

IBGE (2010), Censo demográfico 2010: Características gerais da população, religião e pessoas com deficiência, Relatório técnico, Instituto Brasileiro de Geografia e Estatística. Disponível em: ftp://ftp.ibge.gov.br/Censos/Censo_Demografico_2010/Caracteristicas_Gerais_Religiao_Deficiencia/tab1_3.pdf, Acesso: 02/Ago/2012.

STUMPF, M. R.. **Língua de Sinais**: escrita dos surdos na Internet. V Congresso IberoAmericano de Informática na Educação – RIBIE. Viñadelmar, Chile: [s.n.]. 2000. p. 1-8.

WAUTERS, L. N. (2005), Reading comprehension in deaf children: The impact of the mode of acquisition of word meanings, Tese de doutorado, Radboud University, Nijmegen, Holanda.

WHO (2012), 'Deafness and hearing impairment'. World Health Organization. Disponível em: <http://www.who.int/mediacentre/factsheets/fs300/en/>. Acesso: 02/Ago/2012.