



## **Proposta de Projeto – Fase Piloto**

### **Grupo de Trabalho – Segunda Fase**

**GTAaaS 2.0 – Acessibilidade como um Serviço  
com o foco em pessoas com deficiência visual  
(Segunda Fase)**

Prof. Dr. Tiago Maritan U. de Araújo (UFPB)

Prof. Dr. Guido Lemos Souza Filho (UFPB)

12/Nov/2016

## 1. Visão geral

### 1.1. Descrição do produto/serviço resultante do piloto

Pessoas com deficiência visual possuem dificuldades para captação e compreensão de informações em experiências audiovisuais como vídeos, aulas, cinema, televisão, teatro, óperas, entre outros. Em consequência disso, em muitas situações, elas acabam ficando excluídas do contexto sociocultural e intelectual da sociedade, uma vez que não são disponibilizados recursos de acessibilidade que lhes garantam a participação e o consumo desses conteúdos.

Para reduzir essas barreiras de acesso, um recurso de acessibilidade bastante útil para as pessoas com deficiência visual é a audiodescrição (AD). A AD consiste na descrição de imagens, narração de ações e detalhes visuais como figurinos, indicação de tempo e espaço [9]. Mais especificamente em vídeos, o objetivo da audiodescrição é a tradução dos elementos visuais que é inserida entre os diálogos, através de uma narração, de forma que não haja interferência nos efeitos sonoros do áudio original [1]. Essa tradução transforma toda a informação visual em palavras, ou seja, audiodescreve tudo o que é visto.

Para audiodescrever os vídeos, os roteiros de audiodescrição são geralmente produzidos de forma manual, o que demanda muito tempo. Em média, são necessárias cerca de 60 horas para um profissional descrever um vídeo com duas horas de duração [8], por exemplo. Ademais, segundo pesquisas com empresas que prestam o serviço de audiodescrição, o custo para a criação de um roteiro de AD é de R\$ 40,00 por minuto de vídeo. No entanto, diversos serviços e plataformas de acesso a vídeos ainda não contam com este recurso de acessibilidade, e, em consequência disso, as pessoas com deficiência visual com dificuldade de acesso a este tipo de informação.

Na literatura científica, alguns trabalhos são direcionados para a aplicação da audiodescrição. No caso da audiodescrição específica para vídeos, alguns trabalhos abordam o uso de sintetizadores de voz na narração [2,7,10], a apresentação de diferentes níveis de detalhamento de audiodescrição com versões resumidas e estendidas [2,3] e o uso de players próprios para facilitar a adição e consumo da audiodescrição [2,7]. Quanto à geração automática das descrições, pesquisas, em estágios iniciais, estão sendo feitas com o uso de técnicas de reconhecimento de elementos do vídeo [4] e de análise de roteiro cinematográfico sem vínculos com o tempo do vídeo [8]. Contudo, essas soluções apresentam limitações que ainda não atingem as necessidades de automatização da AD de forma satisfatória. Diante disso, a geração automática ou semi-automática da audiodescrição de vídeos, apresenta-se como uma área pouco explorada.

Para reduzir esses problemas, uma solução automática, ou semi-automática, para gerar trilhas de AD pode auxiliar nesta tarefa, tornando o processo de descrição mais eficiente, diminuindo o tempo investido e o custo associado. Dentro deste contexto, neste projeto, estamos propondo a criação do GTAaaS 2.0 (Acessibilidade como um Serviço - versão 2) que tem como objetivo desenvolver um serviço para geração automática, ou semi-automática, de conteúdos acessíveis para pessoas cegas. A proposta do serviço, denominado Audiodescricao@RNP, é que usuários (ou serviços) submetam vídeos e o serviço piloto gere de forma automática ou semi-automática trilhas de AD, tomando como base o vídeo, o roteiro e as legendas e/ou outros metadados relacionados ao vídeo, caso estejam disponíveis.

Nesta fase do projeto, além de refinar os componentes do protótipo desenvolvido na primeira fase, também serão incorporadas estratégias de reconhecimento de objetos em imagens (por exemplo, *Google Images*<sup>1</sup>) para diminuir a dependência do roteiro, legenda e dos metadados relacionados ao vídeo. Adicionalmente, a ferramenta colaborativa de roteiros de AD, denominada WikiAD, também será evoluída, para permitir que especialistas em AD possam auxiliar na criação destes conteúdos acessíveis. Para isso, a ferramenta será expandida incluindo recursos de *crowdsourcing*, como o gerenciamento de colaboradores, definição de tarefas, ranking de usuários, entre outros.

Com isso, espera-se projetar, desenvolver e ofertar um serviço piloto genérico, com uma API pública e aberta, que permita que os serviços de vídeos da RNP (Cinemas em Rede, Videoaula@RNP, Video@RNP, entre outros) possam utilizá-lo para gerar conteúdos acessíveis para pessoas com deficiência visual.

Uma vez que este serviço piloto possuirá uma API pública e aberta, além de atender e aprimorar os serviços da RNP, ele poderia ser ofertado as instituições clientes da RNP e da administração pública federal, para auxiliá-los no cumprimento dos requisitos da Lei Brasileira de Inclusão (Lei No. 13.146, de 6 de Julho de 2015<sup>2</sup>), que estabelece que estas instituições devem oferecer recursos de legenda oculta, janela com intérpretes de Libras e audiodescrição na promoção de congressos, seminários, oficinas, e demais eventos de natureza científico-cultural:

*“Art. 70. As instituições promotoras de congressos, seminários, oficinas e demais eventos de natureza científico-cultural devem oferecer à pessoa com deficiência, no mínimo, os recursos de tecnologia assistiva previstos no art. 67 desta Lei” (legenda oculta, janela com intérpretes de Libras e audiodescrição”.*

*Art. 71. Os congressos, os seminários, as oficinas e os demais eventos de natureza científico-cultural promovidos ou financiados pelo poder público devem garantir as condições de acessibilidade e os recursos de tecnologia assistiva.*

Adicionalmente, a ANCINE (Agência Nacional de Cinema) instituiu a Instrução Normativa No. 128 de 13 de Setembro de 2016<sup>3</sup>, que estabelece que os todos os filmes exibidos no Brasil sejam distribuídos com recursos de acessibilidade em LIBRAS, legenda descritiva e audiodescrição. Além disso, ela também estabelece um prazo de 24 meses para que todas as salas de exibição comercial disponham de tecnologia assistiva voltada à fruição destes recursos acessíveis (legendagem, legendagem descritiva, audiodescrição e LIBRAS). Neste contexto, o serviço piloto também poderia ser uma alternativa para auxiliar na oferta e atendimento destas demandas.

## **1.2. Identificação do público alvo**

O público alvo deste serviço piloto são as instituições usuárias da RNP que desejem prover conteúdos audiovisuais (vídeos, vídeo-aulas, palestras, eventos científico-culturais, congressos, entre outros) acessíveis para pessoas com deficiência

---

<sup>1</sup> <http://images.google.com>

<sup>2</sup> [http://www.planalto.gov.br/ccivil\\_03/\\_Ato2015-2018/2015/Lei/L13146.htm](http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2015/Lei/L13146.htm)

<sup>3</sup> <http://www.ancine.gov.br/legislacao/instrucoes-normativas-consolidadas/instru-o-normativa-n-128-de-13de-setembro-de-2016>

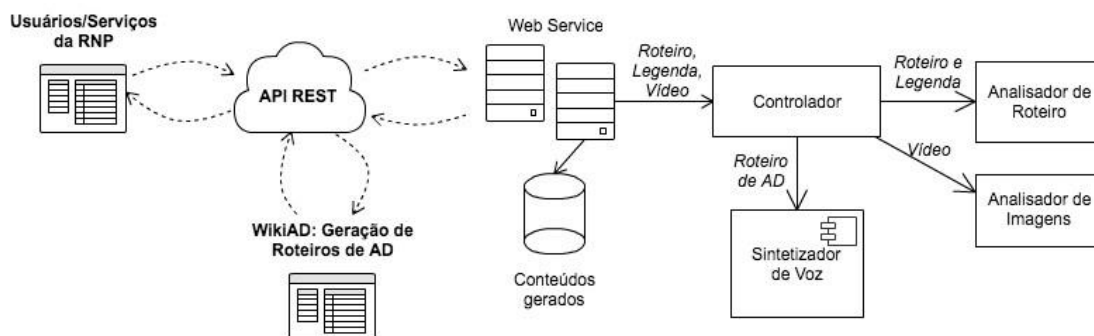
visual, atendendo as exigências da Lei Brasileira de Inclusão, além dos próprios estudantes e usuários com deficiência visual das instituições usuárias da RNP.

É importante destacar que o número de estudantes com deficiência na educação superior vem crescendo bastante nos últimos anos, o que demanda uma maior oferta de tecnologias assistivas e serviços para o atendimento as necessidades desses estudantes. Segundo o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) [6], entre os anos de 2000 e 2015, o número de pessoas com deficiência que se matricularam em cursos superiores presenciais e à distância no Brasil cresceu 1.745%, tendo subido de 2.173 alunos matriculados em 2000, para 37.927 alunos matriculados em 2015.

## 2. Definição do piloto

### 2.1. Arquitetura do piloto

Conforme mencionado na Seção 1.1, nesta segunda fase do projeto, além de refinar os componentes do protótipo do serviço desenvolvido na primeira fase, também serão incluídas estratégias de reconhecimento de objetos em imagens (componente Analisador de Imagens) para reduzir a dependência do roteiro e da legenda, tentando extrair informações para audiodescrição diretamente dos quadros do vídeo. Na Figura 1, é ilustrada a arquitetura deste serviço piloto.



**Figura 1.** Arquitetura do Serviço Piloto (Audiodescrição@RNP)

De acordo com a Figura 1, o serviço piloto funcionará da seguinte forma. Inicialmente, um usuário ou serviço da RNP submete um vídeo para o Audiodescrição@RNP usando uma API REST pública e aberta. Esse vídeo é recebido por um servidor de aplicações que o encaminha para um componente Controlador. Caso o roteiro e a legenda desse vídeo estejam disponíveis, o Controlador envia para o componente Analisador de Roteiro, que identifica os intervalos de silêncio do vídeo, seleciona as sentenças mais importantes a serem descritas, gerando automaticamente um roteiro de AD. Esse roteiro de AD contém uma sequência de descrições e as suas respectivas marcações de tempo. Em seguida, o vídeo é encaminhado para o componente Analisador de Imagens que analisa os quadros do vídeo e extrai informações complementares para a composição do roteiro de AD. Dando continuidade ao processo, o Controlador encaminha o roteiro de AD gerado para o componente de Síntese de Voz, que gera uma trilha de AD sincronizada com o vídeo original. A trilha de AD é então encaminhada/devolvida para o usuário/serviço cliente.

Caso o roteiro e a legenda do vídeo não estejam disponíveis, existirão duas possibilidades. A primeira é que o Controlador encaminhe o vídeo diretamente para o Analisador de Imagens (sem passar pelo componentes Analisador de Roteiro), para gerar uma versão do roteiro de AD a partir do próprio vídeo. A segunda opção é que o vídeo seja encaminhado para a ferramenta WikiAD, uma ferramenta de criação manual de roteiros de AD. Nessa ferramenta, inspirada nos conceitos de computação por humanos e adaptada para geração de roteiros de AD, colaboradores e audiodescritores humanos podem, de forma colaborativa, criar um roteiro de AD para o vídeo. Esse roteiro é então enviado para o Sintetizador de Voz que gerará a trilha de AD para o usuário.

Com relação a ferramenta WikiAD, ela também será evoluída de forma que inclua recursos de crowdsourcing para estimular a participação de colaboradores e especialistas em AD. Dentre as novas funcionalidades previstas estão as seguintes: cadastro e autenticação de usuários, gerenciamento de colaboradores, definição de tarefas/atividades, ranking de usuários, entre outros.

## **2.2. Instituições participantes**

Instituto Benjamin Constant (IBC)

- Contato: a definir
- Descrição: O IBC é uma instituição usuária da RNP e é considerado um Centro de Referência Nacional na Área de Deficiência Visual. O Instituto Benjamin Constant oferece diversos cursos de capacitação na área da deficiência visual, capacita profissionais da área da deficiência visual, assessora escolas e instituições em geral e oferece reabilitação física. Um dos cursos oferecidos pelo IBC é o de audiodescrição.
- Contribuições esperadas para o piloto: Por se tratar de uma instituição de referência na área de deficiência visual, a proposta é que seja executado um piloto do serviço com o IBC e que seus usuários, estudantes e colaboradores possam avaliar o serviço desenvolvido, sugerir melhorias e acompanhar a evolução do serviço piloto.

## **2.3. Objetivos e evoluções**

Nesta fase do projeto, o objetivo geral do projeto é refinar o protótipo e incluir novas funcionalidades como o suporte a estratégias de reconhecimento de objetos em imagens, reduzindo a dependência do roteiro do vídeo para geração dos conteúdos acessíveis. Além disso, pretende-se executar um piloto do serviço com o IBC, uma instituição usuária da RNP que é referência na área de deficiência visual. Para alcançar o objetivo geral foram traçados os seguintes objetivos específicos:

1. Levantamento das principais estratégias de reconhecimento de objetos em imagens disponíveis na literatura científica;
2. Integração e adaptação da estratégia de reconhecimento de imagens ao serviço proposto;
3. Evolução dos componentes de Síntese de Voz para inclusão de novos modelos de voz;
4. Evoluir a ferramenta WikiAD, incluindo recursos de *crowdsourcing* como gerenciamento de comunidade, especificação de atividades/tarefas, ranking de usuários, entre outros;

5. Realizar um piloto com o serviço desenvolvido envolvendo usuários, estudantes e colaboradores do IBC.

### 3. Aproveitamento dos resultados do piloto

Uma vez que o serviço Audiodescrição@RNP dispõe de uma API pública e aberta, o serviço desenvolvido poderá ser facilmente integrado e utilizado por outros serviços e usuários da RNP para geração de trilhas de Audiodescrição para seus conteúdos audiovisuais.

O serviço pode ser disponibilizado pela RNP para a comunidade de duas formas:

- (1) De forma direta: Através de um portal Web em que os usuários/estudantes com deficiência visual submetem de forma direta os seus vídeos e visualizam cópias acessíveis destes vídeos
- (2) De forma indireta, em que os usuários da RNP acessam algum dos serviços de vídeo já ofertados pela RNP (Cinemas em Rede, Videoaula@RNP, Video@RNP) que utilizam o Audiodescriçao@RNP para gerar uma versão acessível dos conteúdos de forma transparente aos usuários finais.

### 4. Macro cronograma de desenvolvimento do piloto

Macro atividades	1º. Trim.	2º. Trim.	3º. Trim.	4º. Trim.
1. Levantamento de técnicas de reconhecimento de objetos em imagens	X	X		
2. Integração/Adaptação de uma estratégia de reconhecimento de imagens no serviço piloto		X	X	X
3. Demonstração dos resultados parciais no WRNP		X		
4. Inclusão de recursos de crowdsourcing no WikiAD	X	X	X	
5. Planejamento do Piloto com o IBC	X	X		
5. Execução do Piloto com o IBC		X	X	X
6. Preparação para o Workshop de Disseminação dos Resultados			X	
7. Apresentação Final dos Resultados			X	
8. Melhorias, refinamentos e finalização da segunda fase do projeto			X	X

## 5. Recursos para o desenvolvimento do piloto

### 5.1. Recursos oferecidos pela RNP para execução do piloto

Considerando que nesta fase do projeto serão incluídas funcionalidades de análise de objetos em vídeos, pode ser necessário a utilização de recursos de processamento distribuído como, por exemplo, os disponíveis no PlanetLab.

### 5.2. Recursos virtualizados para o desenvolvimento do piloto

Nesta fase, a princípio não é necessário que recursos virtualizados sejam disponibilizados pela RNP. No entanto, caso a RNP deseje, pode ser feita uma instalação e instanciação do serviço piloto na infraestrutura da RNP.

## 6. Cronograma e entregas pré-definidas

### 27/01/2017

- RP4: Planejamento de Recursos Virtualizados
- RP5: Planejamento da estrutura de pacotes de trabalho de desenvolvimento tecnológico e do cronograma de entregas destes pacotes

### 24/02/2017

- RT4: Relatório de mapeamento de componentes e licenças de software
- RT5: Plano de testes do piloto **31/03/2017**
- Site do GT atualizado
- Iniciar a implantação do piloto<sup>4</sup>
- RP6: Relatório de planejamento do WRNP (demonstração, material e viagens)
- RA5: Relatório de acompanhamento trimestral jan/fev/mar **28/04/2017**
- Entrega do **código-fonte** da versão implantada no piloto (códigos-fonte, executáveis, *scripts*, arquivos de configuração etc.), incluindo o sistema e as ferramentas de suporte à operação;
- Entrega de **documentação** (manuais de instalação e administração, manuais de usuário etc.).

### 15/05/2017 a 16/05/2017

- WRNP: Apresentação em sessão técnica e demonstração dos resultados parciais do piloto no Workshop RNP nos dias 15 e 16/05 em Belém, PA. **30/06/2017**

---

<sup>4</sup>Início das atividades planejadas no RP6.

- RWRNP: Relatório de participação no WRNP
- RA6: Relatório de acompanhamento trimestral abr/mai/jun **28/07/2017**
- RP7: Relatório de planejamento do Workshop de Disseminação do GT

**25/08/2017**

- RT6: Avaliação dos Resultados do Piloto
- RT7: Recomendações para a implantação do serviço/produto

**Entre 01/09/2017 a 30/09/2017 (data a definir)**

- Realização do Workshop de Disseminação do GT (data a definir)

**Entre 01/10/2017 a 31/10/2017 (data a definir)**

- Apresentação Final dos Resultados para o comitê de avaliação
- RA7: Relatório de acompanhamento trimestral ago/set/out **24/11/2017**
- RP8: Relatório de planejamento para inclusão no portfólio da RNP
- Atualização do RT4: Relatório de mapeamento de componentes e licenças de software
- Entrega final do **código-fonte e documentação 15/12/2017**
- RA8: Relatório de acompanhamento out/nov/dez

## 7. Referências

[1] BENECKE, B. **Audio-description**. In: GAMBIER, Y. (ed.), 2004. Meta. Volume 49, nº.1, p. 78-80.

[2] CHAPDELAINE, C.; GAGNON, L. **Accessible Videodescription On-Demand**. ASSETS '09, Pennsylvania, p 221-222, oct. 2009.

[3] ENCELLE, B.; BELDAME, M. O.; PRIÉ, Y. **Towards the usage of pauses in audio described videos**. 22nd International World Wide Web Conference, Rio de Janeiro, may 2013.

[4] GAGNON, Langis et al. **Towards computer-vision software tools to increase production and accessibility of video description for people with vision loss**. Springer-Verlag, p 199-218, feb. 2009.

[5] IBGE. Censo Demográfico 2010: Características gerais da população, religião e pessoas com deficiência. **IBGE- Instituto Brasileiro de Geografia e Estatística**, 2010. Disponível em: <<http://bit.ly/1S46pl4>>. Acesso em: 28/01/2014.

[6] INEP. Sinopses Estatísticas da Educação Superior – Graduação. **INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira**. Disponível em: <<http://portal.inep.gov.br/superior-censosuperior-sinopse>>. Acesso em: 12/12/2016.



- [7] KOBAYASHI, M.; NAGANO, T.; FUKUDA, K.; TAKAGI, H. **Describing Online Videos with Text-to-Speech Narration**. W4A2010, Raleigh, apr. 2010
- [8] LAKRITZ, J.; SALWAY, A. **“The Semi-Automatic Generation of Audio Description from Screenplays”**, Dept. Of Computing Technical Report CS-06-05, 2002, University of Surrey.
- [9] NUNES, E. V.; MACHADO, F. O. e VANZIN, T. (2011). **Audiodescrição como Tecnologia Assistiva para o Acesso ao Conhecimento por Pessoas Cegas**. Ambiente virtual de aprendizagem inclusivo / organizadores: Vania Ribas Ulbricht, Tarcísio Vanzin e Vilma Villarouco. – 352 p. Florianópolis : Pandion, 2011.
- [10] O'CONNELL, T.; KOBAYASHI, M.; GOULD, B.; TAKAGI, H.; ASAKAWA, C. **Are Synthesized Video Descriptions Acceptable?** ASSETS '10, Florida, p. 163-170, oct. 2010.