IBM® AbilityLab™
Media Captioner and Editor©
IBM AbilityLab Media Captioner & Editor (MCE)

Speaker: Ali Sobhi (sobhi@us.ibm.com)

Authors: Ali Sobhi, Reiko Nagatsuma

Session ID: DHH-014

Wednesday February 29, 2012
10:40 AM – 11:40 AM

Grand Hyatt – Elizabeth F
IBM Media Library Metrics

Total media count = 78,546
Media w/o transcript = 93%

A great portion of all media content on Media Library does not have captions or transcriptions

IBM Media Library

Media & Transcripts - June 2010

- MP3: 69.79%
- MP3+transcript: 23.27%
- Video: 4.70%
- Video+transcript: 2.25%
## Matrix of Media and Technology Components

<table>
<thead>
<tr>
<th>Media Source</th>
<th>Digital Recording</th>
<th>Media Transformation</th>
<th>Transcription</th>
<th>Edit</th>
<th>Translation</th>
<th>Edit</th>
<th>Media Generation / Synchronization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broadcast - TV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Live</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Recorded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broadcast – Radio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Webcast</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Podcast</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Audio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Video</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture Presentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Speaker notes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Charts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VoIP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Teleconference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• One-to-one</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• IM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• A / V Chat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Media Captioner & Editor Technology Effort

Corporate requirement to caption videos published on www.ibm.com and w3

Provide Captioning, Transcription and Editing Capability for Ever Increasing Multimedia Content On IBM w3 Intranet

A Collaborative Effort between IBM Research Groups

Human Ability & Accessibility Center China Research Lab Tokyo Research Lab

Using state-of-the-art IBM Automatic Speech Recognition
Internal Announcement of MCE

Technology Adoption Program

IBM AbilityLab Digital Multimedia Captioner & Editor - (DigiCapE)

Web-based multimedia captioning and multi-user editing - DigiCapE

Downloads: 213
Created: 01 Nov 2010
Updated: 01 Nov 2010
Innovators: (View all Innovators)

Ali Sobhi
Solution Architect - Human Ability & Accessibility Center, Team Lead

Michael Picheny
Manager, Watson Research

IBM © 2012 IBM Corporation

Feb 29, 2012

IBM Media Captioner & Editor
MCE Users

**HA & AC Deploys Video Captioning and Editing System (Nov 2008)**

- Embraced by W3 team to caption executive videos
- Alternative to current expensive commercial captioning service
- Result is a more cost effective and faster solution

**What W3 Team is Saying**

- “Love this tool ... was able to transcribe/caption a 4 minute video in under an hour. This could really change how the w3 team does a lot of things …”
- “I’m really, really happy with the results. Really happy. Thanks for a great tool.”
- “I'm getting more and more people up and running on DigiCapE …”

**The Project**

- World-wide collaboration between HA & AC, CRL, TRL and WRL
- Speaker independent Automatic Speech Recognition WRL
- Caption Editing Client (**CES**) – TRL
- Closed-Caption Media (**BCC Platform**) – CRL
- API design, communication architecture and deployment – HA&AC
CIO first quarter update
Mark Hennessy on First Quarter Performance
Captioned video

hello to those of you in Paris.

Kelly: Let's begin with kind of an

Watch the interview, become a cloud insider and
start playing in the sandbox.

Visit Cloud on TAP
IBM AbilityLab Media Captioner & Editor

- Complete solution to caption, edit and create Web-ready digital media
- Helping people who are deaf or hard of hearing
- Utilizing state-of-the-art advanced, speaker-independent IBM Automatic Speech Recognition

Intuitive and easy-to-use caption editor with synchronized audio and editing

Web-based GUI for complete job management

Web-ready media with closed-captioning enabled
Example of MCE Captions

HR leader for the People with Disabilities Global Diversity Council.

Líder de RH del Consejo de Diversidad Global para individuos incapacitados.
MCE Jobs and Captioned Media Stats – July 2010

DigiCapE Jobs & Captioned Media Hours
To-Date: July 2010

- Total Number of Jobs = 2100
- Total Media Hours = 596
MCE Jobs and Captioned Media Stats – July 2011

DigiCapE Monthly Jobs & Media (Hours)
To-Date: July 2011

Total Number of Jobs = 3680
Total Media Hours = 1005
Objectives:
- Video captioning based on speech recognition with editing capability
- Available on SCIC for pilots – 2012

Status:
- Internal Use
- Available on SCIC for pilots – 2012
Usage Scenario

Who will use this?
- PwD – deaf / hard of hearing
- Captioning of media archives
- Search engines
- Data mining

Examples:
- IBM
- NCSU
- NASA
- National Archives
- Broadcasting companies

How to use this?
- Login to the Web
- Upload appropriate media format to server
- Notification e-mailed - “Caption Ready”
- If still logged in, status gets updated on Web GUI
- Edit the captions, as needed, synchronized with audio track
- Submit for final merge – Open caption / Closed caption
- Show-off your handiwork and get bonuses
Possible Business Model

Client & Web Services Model

Client:
- Direct sale
- Updates
- Upgrades
- Training

SOA Services:
- Services-based
- Registration fee
- Usage fee
- Can be upgraded w/o any service interruption

Enterprise Solution Model

Client:

Server

Services

Transcription Text

HTML SMIL

Enterprise Sale
Complete sale and service inside
Transcribed with IBM DigiCapE

where I bring you information on new and emerging technologies from IBM.
Harness the power of speech recognition to create high-quality captioned content

Highlights

- Uses advanced automatic speech recognition technology to analyze and "translate" spoken content to a text transcript.
- Offers a full-screen editor with synchronized text and voice playback for precise correction of the transcript.
- Creates final media content with high-quality captions by combining the original media content with the corrected caption.

From bare to there, in a few hours

If you have a bare Webcast, podcast, or digital video or audio, in need of captions, you can use the IBM® AbilityLab Media Captioner and Editor to create them automatically!

The IBM AbilityLab Media Captioner and Editor has two parts:

- The media captioning server – this Web-based program accepts files to be captioned. The server reads your submitted file several times, using advanced automatic speech recognition technology to interpret the spoken words and create a "written" transcript. After a period of time – several minutes to overnight, depending on the size of your content file – the media captioning server creates and sends the written transcript to your system for you to review, edit, and approve.
IBM AbilityLab Media Captioner and Editor

Extend the impact of your digital multimedia content with automated, advanced captioning from IBM

Generate value with captioned:

- Videos
- Webcasts
- Podcasts

Captioning and transcription allow you to expose your rich digital content to a broader audience, including people who are deaf, hard of hearing and non-native language speakers who benefit from multimodal content delivery. This versatile technology can also help simplify digital archiving processes and enhance search engine optimization.

IBM AbilityLab Media Captioner and Editor makes the process of adding captions easier than ever, with an automated solution that gives you maximum flexibility and control to quickly generate accurate captions and transcripts for videos, podcasts, webcasts and other multimedia.

Simply upload your digital content by clicking the Caption Now button to the right to receive an automatically generated draft transcript of your multimedia file. Using the AbilityLab Media Captioner & Editor interface, you can then edit the transcript for accuracy and create a captioned version of your original digital file. The resulting transcription files can then be used with popular online video players like YouTube.

Note: AbilityLab Media Captioner and Editor is an IBM Research asset. It is not yet a formal IBM product offering. Select clients and partners are approved to test AbilityLab Media Captioner and Editor in an IBM pre-production cloud environment. IBM-approved login credentials are required for access. To explore the possibility of piloting this Research technology in your organization, please send an email to research@us.ibm.com.
IBM AbilityLab Media Captioner & Editor (MCE)

MCE allows users to upload or link to video, create and edit transcripts from the video, then create a captioned version of the video. MCE also creates transcription files which can be used with other video players (YouTube, for example).

Notice - updated Wednesday February 22, 2012
This host is only for test and development of MCE
Version: 2.1.48-2012-02-20
Release Type: Proxy Release

Scheduled Maintenance: every Sunday 8:00 AM - 8:00PM CST
Please provide the following information with your communications:
JobID, MCE version, browser used and its version, OS Platform
Thank you for your continued support of MCE

MCE works best with Firefox 8 or greater
Features Rolling out on SCIC

- **MCE** is available in “pilot” capacity on **SmartCloud Innovation Center (SCIC)**
  - Contact MCE Web site for information on how to participate
  - Features will be activated as they are certified

- User community and management
  - Allow multiple people share the editing tasks
  - Who edited which line?
  - Change ownership

- Dojo 1.6 keyboard navigation and management

- Closed-caption

- Creating sample HTML drop-in code-segment
  - ccPlayerAS3.swf FlashPlayer (WGBH), swobject.js (Google)

- Creating other caption file format
  - SMIL, RealText, QuickText

- Post-ASR processing
  - Analytics, Edit-assist, feed back to language domain improvements

- Additional ASR languages models and domain