



# IBM® AbilityLab™ Digital Media Captioner & Editor

# IBM AbilityLab Digital Media Captioner and Editor

Speaker: **Ali Sobhi**

[sobhi@us.ibm.com](mailto:sobhi@us.ibm.com)

Session ID: DHH-2003

Wednesday March 16, 2011

10:40 AM – 11:40 AM

Grand Hyatt – America's Cup C

## IBM Media Library Metrics

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

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

# Matrix of Media and Technology Components

This is complex – there are many components – such as the media itself:

- Broadcast - TV
  - Live
  - Recorded
- Broadcast – Radio
- Web
  - Webcast
  - Podcast
  - Audio
  - Video
- Lecture Presentation
  - Speaker notes
  - Charts
- VoIP
  - Teleconference
  - One-to-one
- Internet
  - IM
  - A / V Chat

And the process:

1. Digital Recording
2. Media Transformation
3. Transcription
4. Edit Translation
5. Edit Media Generation / Synchronization

# Media Captioning & Editing Technology Effort

Corporate requirement to caption videos published on [www.ibm.com](http://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

# Captioner and Editor 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 Captioner & Editor ...”

### 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



# Used to caption IBM Executive videos for www and w3

**CIO first quarter**  
**Mark Hennessy on First**  
**Caption**

Corporate Headquarters  
Armonk, NY

hello to those

Instantly deploy to the cloud

Kelly: Let's begin with kind of an

(26 MB, .mp4, please wait while the video loads)

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

[Visit Cloud on TAP](#)

The **Cloud Sandbox** is IBM's first cloud computing environment. It offers our developers a self-service, dynamic computing capacity environment. A select group of testers from the IBM worldwide cloud lab began putting the pilot through the motions in February. The pilot, now on TAP, is ready for a larger audience.

Using the Cloud Sandbox, IBM developers can provision or de-provision a server or group of servers in minutes, as computing requirements change. The sandbox provides a browser-based management user interface, as well as a fully functional [RESTful](#) web services interface.

Like most TAP projects, usage is free—the only cost is the willingness to give the cloud a workout and provide feedback to the development team. The project team is refining and prioritizing new features based on feedback from the community and has big plans to quickly turn this into a client offering later in the year.

**Know any candidates?**

This type of cloud is something most CEOs won't even know they're using — things just, well, work. But for some CEOs, hearing matters. Who are those people?

**Meet Mark Hennessy**

Mark Hennessy has the smarts to match his charm. He learned his latest innovation after it. Hennessy and his team are now enhancing [CloudBurst](#), an OpenSource Internet-scale SaaS accessible database, to provide a database cloud service. They are using cloud sandbox servers to test scalability and multi-tenancy of their system, ensuring that it is optimized for massive data throughput.

Hennessy said, "Within 20 minutes and almost no effort, I was able to create and get full access to an internally accessible Linux image for testing CloudB. Doing the same without the service can take weeks."

Know any CEOs like Hennessy? Send the developers you know an [IBM Cloud Sandbox a card](#) to encourage them to join IBM's efforts to make cloud a reality for your favorite IBM tools (Media Library, Callout ... dare we suggest, the next pilot version of Lotus Notes).

# Captioner and Editor is Easy to Use



IBM Caption Editing System Master

	Start Time	State	Choice	Caption
89	00:02:42.519	<input type="radio"/>	<input type="radio"/>	to the line teams.
90	00:02:44.220	<input type="radio"/>	<input type="radio"/>	Well, said another way
91	00:02:45.619	<input type="radio"/>	<input type="radio"/>	we are trusted process advisors
92	00:02:47.879	<input type="radio"/>	<input type="radio"/>	which
93	00:02:48.089	<input type="radio"/>	<input type="radio"/>	as
94	00:02:48.669	<input type="radio"/>	<input type="radio"/>	the overriding principle
95	00:02:49.979	<input type="radio"/>	<input type="radio"/>	on which finance builds
96	00:02:51.119	<input type="radio"/>	<input type="radio"/>	our
97	00:02:51.489	<input type="radio"/>	<input type="radio"/>	eyes.
98	00:02:52.049	<input type="radio"/>	<input type="radio"/>	Terrific. And
99	00:02:52.889	<input type="radio"/>	<input type="radio"/>	Carlos you mentioned our values.
100	00:02:54.690	<input type="radio"/>	<input type="radio"/>	What are the IBM values mean to Finance organiz
101	00:02:57.679	<input type="radio"/>	<input type="radio"/>	Well Nancy, first and foremost
102	00:02:59.639	<input type="radio"/>	<input type="radio"/>	finances civil use read an organization
103	00:03:02.690	<input type="radio"/>	<input type="radio"/>	and to consultate our values
104	00:03:04.229	<input type="radio"/>	<input type="radio"/>	into principles for our finance people
105	00:03:06.869	<input type="radio"/>	<input type="radio"/>	well
106	00:03:07.809	<input type="radio"/>	<input type="radio"/>	we ask our
107	00:03:08.800	<input type="radio"/>	<input type="radio"/>	managers
108	00:03:09.979	<input type="radio"/>	<input type="radio"/>	fifteen hundred of them worldwide
109	00:03:11.929	<input type="radio"/>	<input type="radio"/>	led the values meant to finance
110	00:03:14.099	<input type="radio"/>	<input type="radio"/>	they told us that we understand who our custome
111	00:03:17.449	<input type="radio"/>	<input type="radio"/>	both internal
112	00:03:19.100	<input type="radio"/>	<input type="radio"/>	like a lot of management team
113	00:03:21.040	<input type="radio"/>	<input type="radio"/>	an external
114	00:03:21.930	<input type="radio"/>	<input type="radio"/>	like IBM customers

Entire time: 00:10:50.988  
Actually play time: 00:10:50.988  
Max characters per line: 60

Before 1 Lines  
After 1 Lines

File Name: C:\Test\finproc\Edited\workspace\2009\_04\_21\_How\_ Elapsed Time: 00:00:23



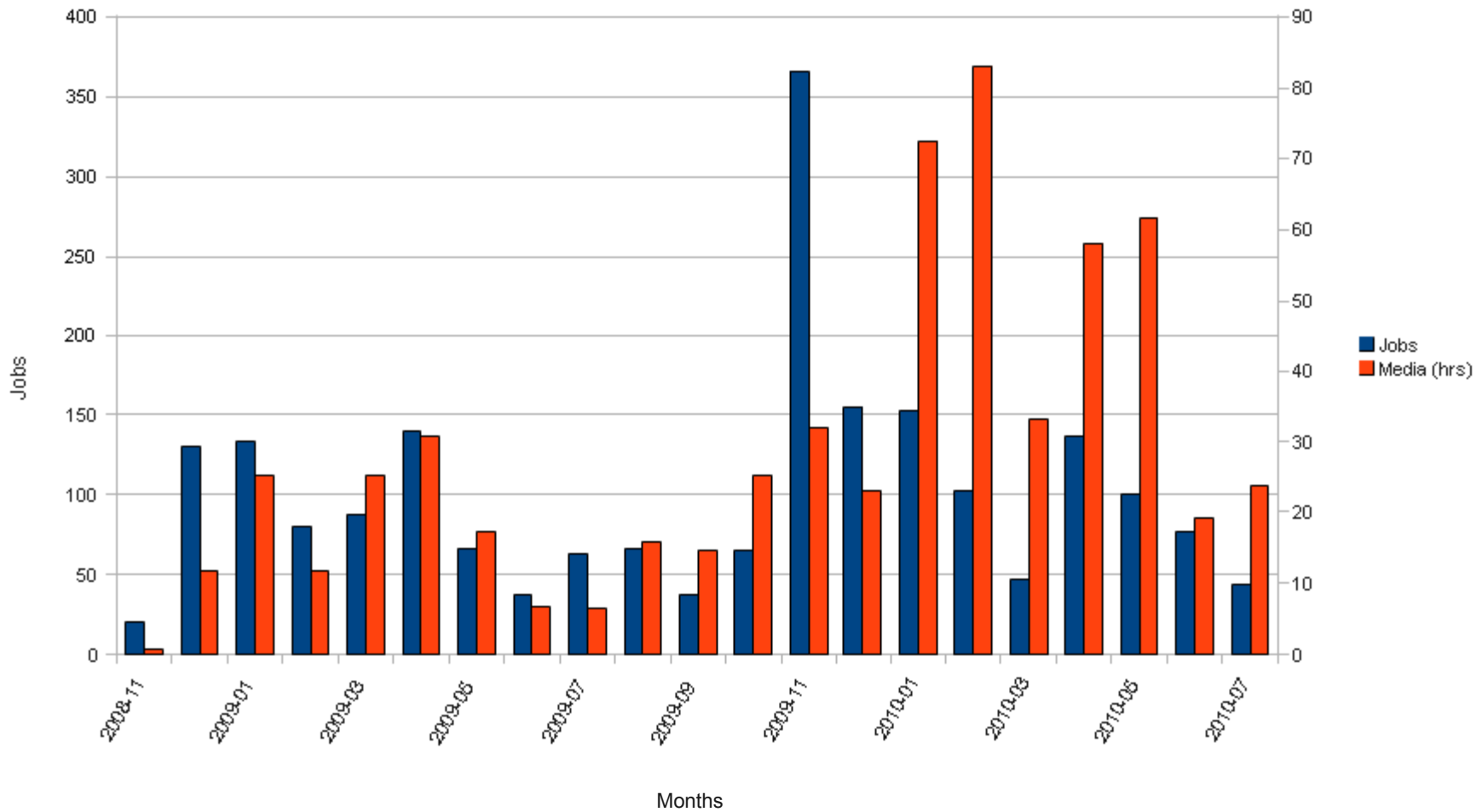
Speech Recognition

Caption Correction

Merged Media Creation

# Captioner & Editor Jobs and Captioned Media Stats

Captioner & Editor Jobs & Captioned Media Hours - To-Date: July 2010



Total Jobs: 2100 (+44)  
 Total hours: 596 (+23)

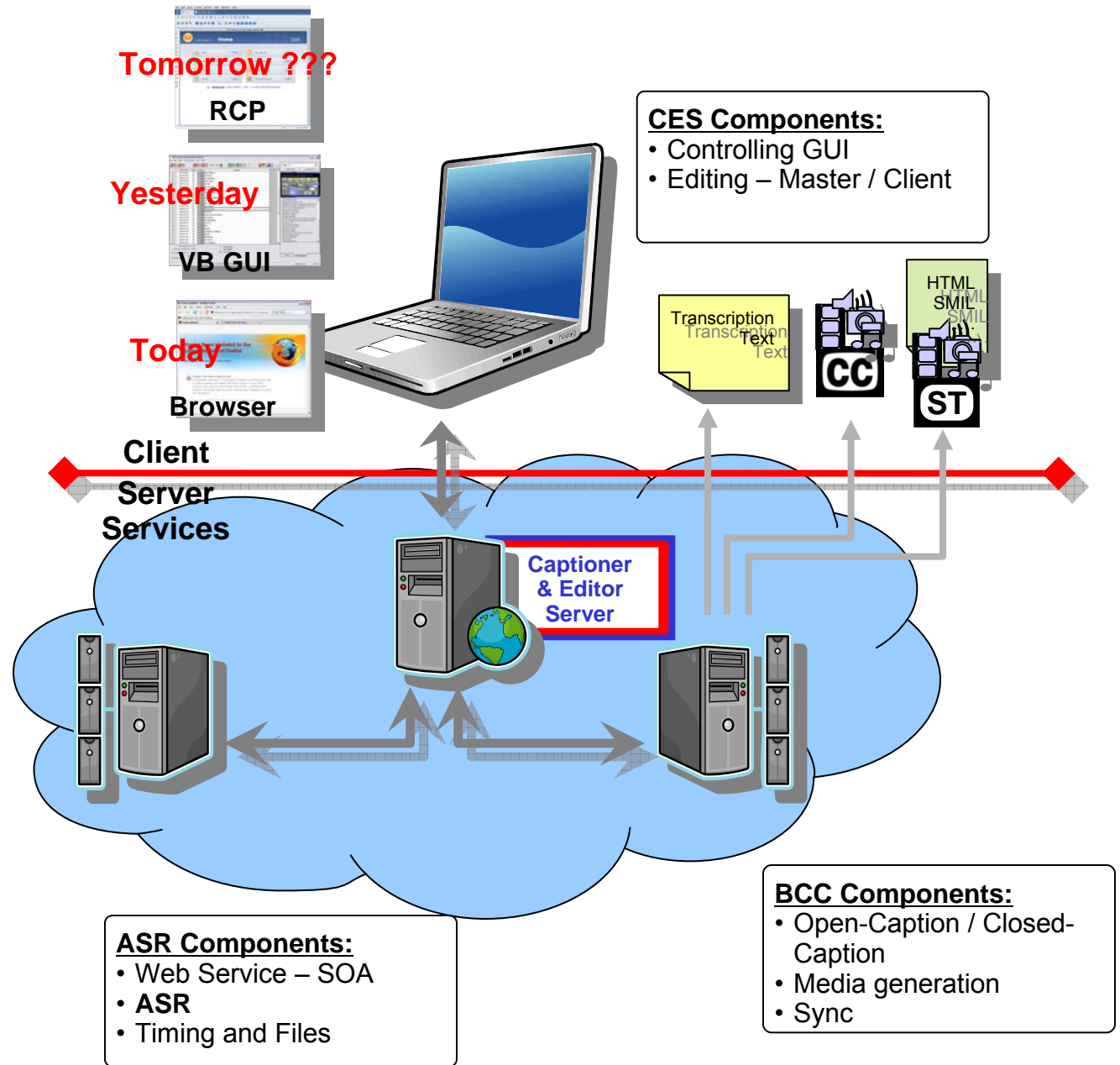
# System Overview and Architecture

## Objectives:

- Video captioning based on speech recognition with editing capability

## Status:

- Internal Use
- **Available on IME for external users – 2011 (March - April)**



# 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 handy work and get bonuses



# Media Captioner & Editor on YouTube

**You Tube**  Search | [Brows](#)

This video is unlisted. Only those with the link can see it. [Learn more](#)

 **Transcribed with IBM DigiCapE**

[katadada](#) 2 videos



NC STATE UNIVERSITY

IBM

where i bring you information on new  
Emerging Technologies  
and emerging technologies from i. b. m.

0:12 / 7:20 CC 360p

# Media Captioner & Editor on YouTube

The image shows a YouTube video player interface. At the top left is the YouTube logo. To its right is a search bar with the word "Search" and a "Brow" link. Below the search bar is the channel name "IBM AbilityLab Media Captioner and Editor" with a dropdown arrow. Underneath the channel name, it says "IBMAccessibility" with a "5 videos" dropdown and a "Subscribe" button. The video player itself shows a slide with the IBM logo in the top right corner, the text "IBM Human Ability and Accessibility Center", and the title "IBM AbilityLab Media Captioner and Editor Demonstration". Below the title, it says "September 2010". The slide features four images: a man in a wheelchair, a woman wearing a headset, a hand holding a document, and a close-up of an ear. At the bottom of the slide, there is a copyright notice "© 2010 IBM Corporation" and a subtitle that reads "Hello everyone and welcome to this demonstration of the IBM AbilityLab Media Captioner and". The video player controls at the bottom show a play button, a volume icon, a progress bar at "0:05 / 4:27", a Creative Commons license icon, a resolution of "360p", and other standard video controls.

# IBM Research Labs Experimental Technology Site

**IBM** United States [change]

Home Solutions Services Products Support & downloads My IBM Welcome Mr. Ali Sobhi [Not you?] [IBM Sign in]

**IBM Research**

- IBM Research Labs Experimental Technology Site
- About
- FAQ

## IBM Research Labs

Welcome to the IBM Research Labs Experimental Technology Site.

Come play with our latest research experimental technologies. We invite you to experiment with these and future beta technologies with us. Your feedback will be essential - we look forward to collaborating with you!

**Feedback**

- Feedback Forum

**IBM Easy Web Browsing**

Helps people with limited vision and seniors access Web content more easily.

- Browse this site with EasyWeb
- Learn more about EasyWeb

**Search Research**

Browse people

Browse projects

Browse papers

**Experiments** Client Direct

Select one

1-5 of 11 experiments | Previous Next Results per page: 5 | all

- IBM Mobile Systems Remote**  
Manage IBM Servers with your Mobile Smartphone  
04 Mar 2011 | ★★★★★ 1 Ratings | View Details
- IBM AbilityLab Voice Transcription Manager**  
Manages and distributes real time multi-user transcriptions of conversations using speech recognition  
07 Dec 2010 | ★★★★★ 0 Ratings | View Details
- IBM AbilityLab HTML Accessibility Checker**  
Make your HTML accessible  
05 Nov 2010 | ★★★★★ 1 Ratings | View Details
- IBM AbilityLab Virtual Worlds Accessible U/I**  
Enable virtual worlds to be accessible for blind users.  
05 Nov 2010 | ★★★★★ 1 Ratings | View Details

<http://www.researchlabs.ibm.com/portal/>

## Features Rolling out on Research IME

- ▶ Look for Media Captioner & Editor on IME
  - March – April
  - Features will be activated as they are certified
- ▶ User community
  - Allow multiple people share the editing tasks
  - Who edited which line?
- ▶ Closed-caption
- ▶ Creating other caption file format
- ▶ Creating sample HTML drop-in code-segment
  - ccPlayerAS3.swf FlashPlayer (WGBH), swobject.js (Google)
- ▶ Post-ASR processing
- ▶ Additional ASR languages