



Session ID: DHH-1010
Wednesday, March 18, 2009
3:10 PM – 3:40 PM
Marriott – Saddle Brook Rm

Speaker:
Ali Sobhi



Agenda

IBM Digital Multimedia Captioning and Editing System

Caption Editing System

Demo



Objectives & Deliverables

Define an appropriate captioning / transcribing and editing solution for HA & AC

- What we think is appropriate and makes sense

- What we help deliver through existing components and new work

Deliver a captioning system that can process media in the IBM media library

- Align with CIO requirements for accessible media

Create asset that can be commercialized

- Assist HA & AC service delivery organization

Digital Captioning and Editing



Matrix of Media and Technology Components

Media Source	Digital Recording	Media Transformation	Transcription	Edit	Translation	Edit	Media Generation / Synchronization
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							

Complex
Many components



Products / Components Consideration

Caption Editing System (CES from TRL)
Editing (master / client), Controlling GUI



Broadcast Captioning (BCC from CRL)
Closed-caption merge and media generation



HLT's new ASR engine controller
API control for Automatic Speech Recognition (ASR) web service





High-Level Architecture



← Future

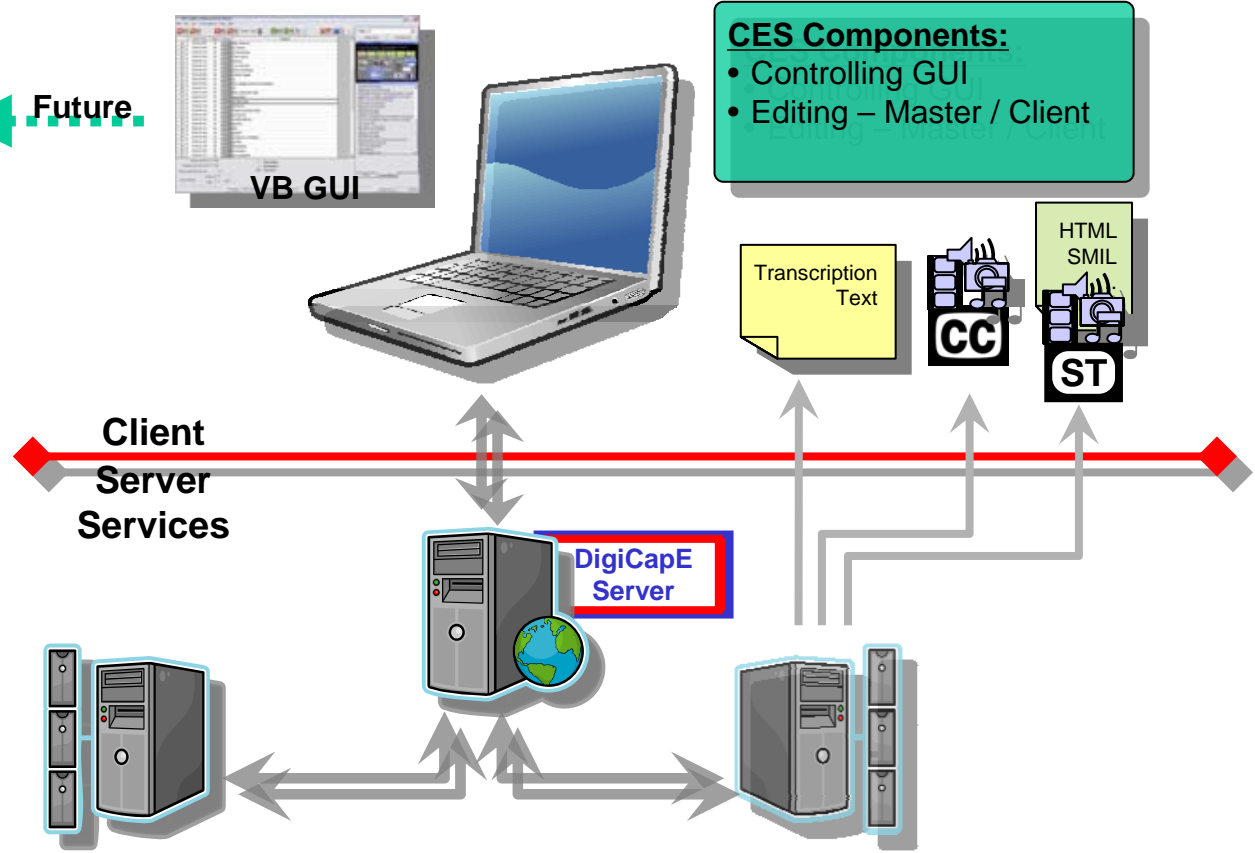


CES Components:

- Controlling GUI
- Editing – Master / Client

Requirements:

- Media upload
- Multi-user editing
- Speaker-independent
- High accuracy
- Batch processing
- Closed-captioned media
- Web-based & SMIL
- Transcription text



ASR Components:

- Web Service – SOA
- ASR
- Timing and Files

BCC Components:

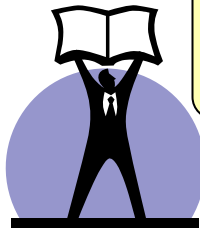
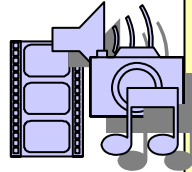
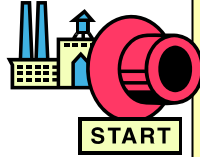
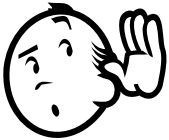
- Closed-Caption
- Media generation
- Sync & highlighting



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?

- Install / configure client GUI
 - Directory and file structure for I/O
 - Define the DigiCapE server URL
- Upload **appropriate media format** to server
- Client GUI will be notified when the transcription is complete
 - The transcript with timing marks is sent back to the client and stored on local disk
- The user can then replay the media and view the transcript in the editing pane
 - User can edit the transcripts and adjust the speed and characteristics of the transcription
- The user uploads the modified transcription bundle back to DigiCapE server
 - Request type of output; Web, single media file
- The user is notified of the completion and the requested output media is stored locally on client



Possible Business Models

Client & Web Services Model

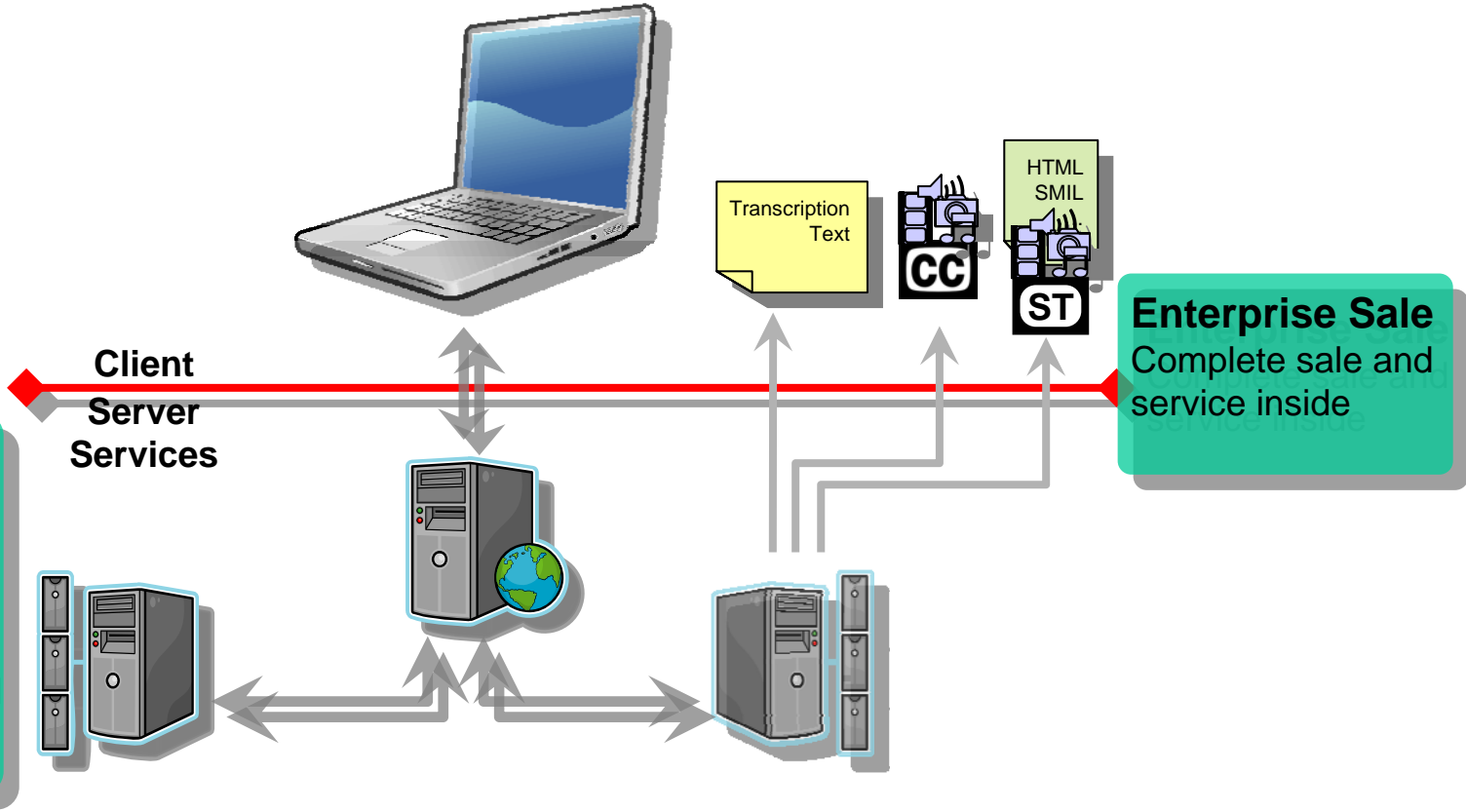
Enterprise Solution Model

Client:

- Direct sale
- Updates
- Upgrades
- Training

SOA Services:

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





Agenda

IBM Digital Multimedia Captioning and Editing System

Caption Editing System

Demo



Caption Editing System: Overview

Captioning for Digital Archive of Lectures, Videos on the Web

Social Merit of Captioning

Accessibility needs for Seniors, Deaf or Hard of Hearing, Foreigners

Regulations and Guidelines

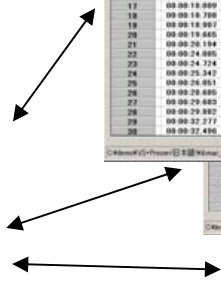
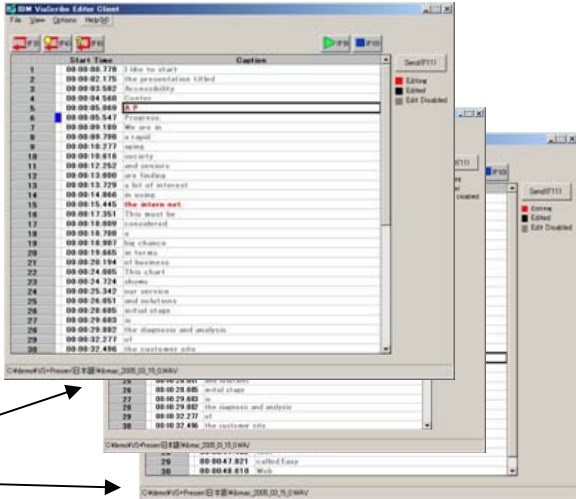
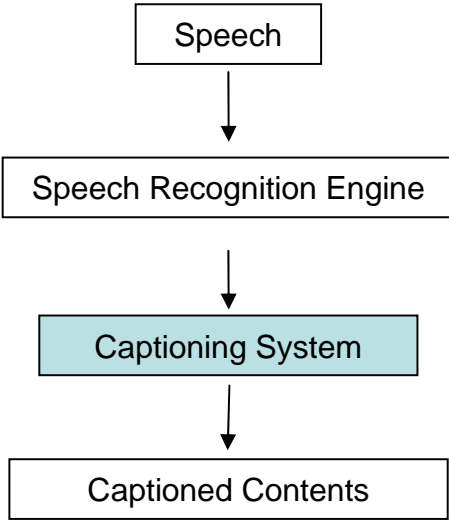
Section 508, WCAG 2.0, National Standards, Cooperate Instructions



There are still many technical issues to captioning

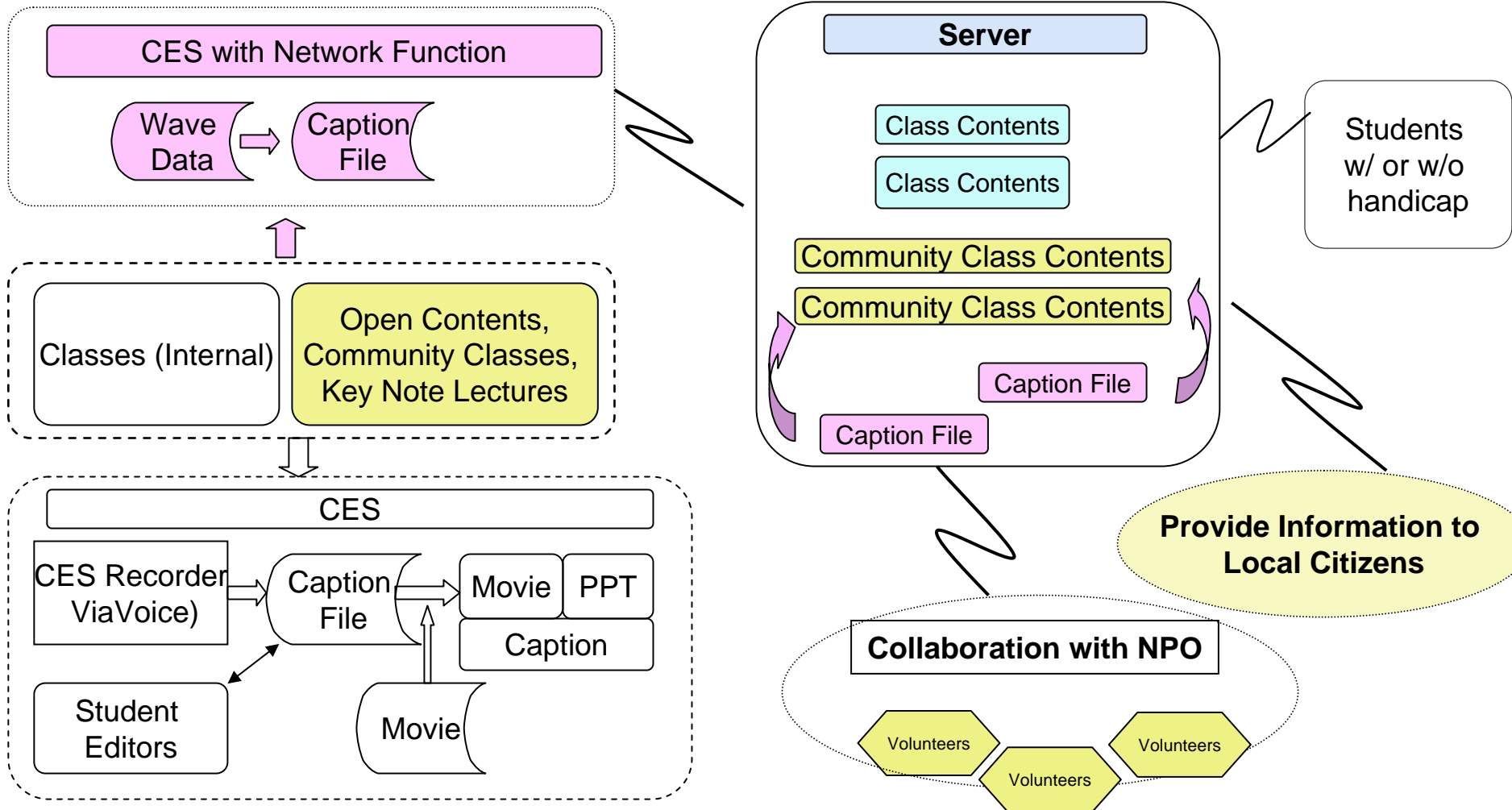
Client Caption Editing System

Master Caption Editing System





The Big Picture: Collaborative Editing Community

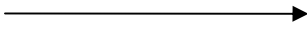




Collaborative Editing by Role

One Skillful and Responsible Master Editor just flags right (O) or wrong(?) or delete (X) by listening to the audio

Novice Client Editors makes the actual correction

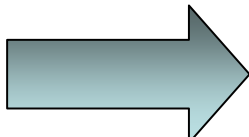


IBM ViaScribe Editor Master

Start Time	State	Chocic
00:00:36.160	?	bond to
00:00:35.679	O	improvement plan
00:00:37.485	O	and then
00:00:37.024	O	verifications
00:00:38.971	O	and in
00:00:39.311	O	the
00:00:39.460	?	lawyer
00:00:39.939	O	prepare a design
00:00:41.027	O	clude
00:00:41.396	O	template.
00:00:43.002	O	we have
00:00:43.431	O	an analysis
00:00:44.160	O	tool
00:00:44.579	O	called
00:00:44.928	?	is a designer
00:00:46.474	O	and Enhancement
00:00:47.402	O	tool
00:00:47.821	O	called Easy
00:00:48.610	O	Web
00:00:48.909	O	Browsing.
00:00:53.199	O	As
00:00:53.469	O	For deployment
00:00:54.227	O	in Japan
00:00:55.195	O	as you can see
00:00:56.242	O	me
00:00:56.452	O	have quite
00:00:57.090	O	a number
00:00:57.539	O	of wins
00:00:58.347	O	with easy
00:00:58.896	O	web browsing.
00:01:03.107	O	Solution highlights

Number of waiting Clients: 0
Number of waiting messages: 0

FTP or LAN



IBM ViaScribe Editor Client


Start Time	Caution
00:00:00.770	I like to start
00:00:02.175	the presentation titled
00:00:03.582	Accessibility
00:00:04.550	Center
00:00:05.009	A P
00:00:05.547	Progress.
00:00:09.109	We are in
00:00:09.798	a road
00:00:10.277	aged
00:00:10.616	society
00:00:12.252	and sensors
00:00:13.000	are finding
00:00:13.729	a lot of interest
00:00:14.066	in using
00:00:15.445	the intern net.
00:00:17.051	This must be
00:00:18.000	considered
00:00:18.708	a
00:00:18.907	big chance
00:00:19.605	in terms
00:00:20.194	of business
00:00:24.005	This chart
00:00:24.724	shows
00:00:25.342	our service
00:00:26.051	and solutions
00:00:28.605	initial stage
00:00:29.603	is
00:00:29.982	the diagnosis and analysis
00:00:32.717	of
00:00:32.496	the customer site

Aligned buttons allow easy flag ops


- Complete Audio Synchronization
- Target Line Emphasis Audio Play
- Mouse-free Operation
- Automatic Jump




Complete Audio Synchronization

 Ondemand Business

5	2.715	Onde And
6	3.393	Business



<Initially, Focus on the Audio>

 **I got it!!**
<Audio Stops>

5	2.715	Onde And
6	3.393	Business


<Key Input>
Cursor to the right
Delete


???

“Onde ...”
What was it?

<Key Input Stops>

<Audio restarts automatically>

 Ondemand Business



<Audio Stops>

5	2.715	Ondemand
6	3.393	Business

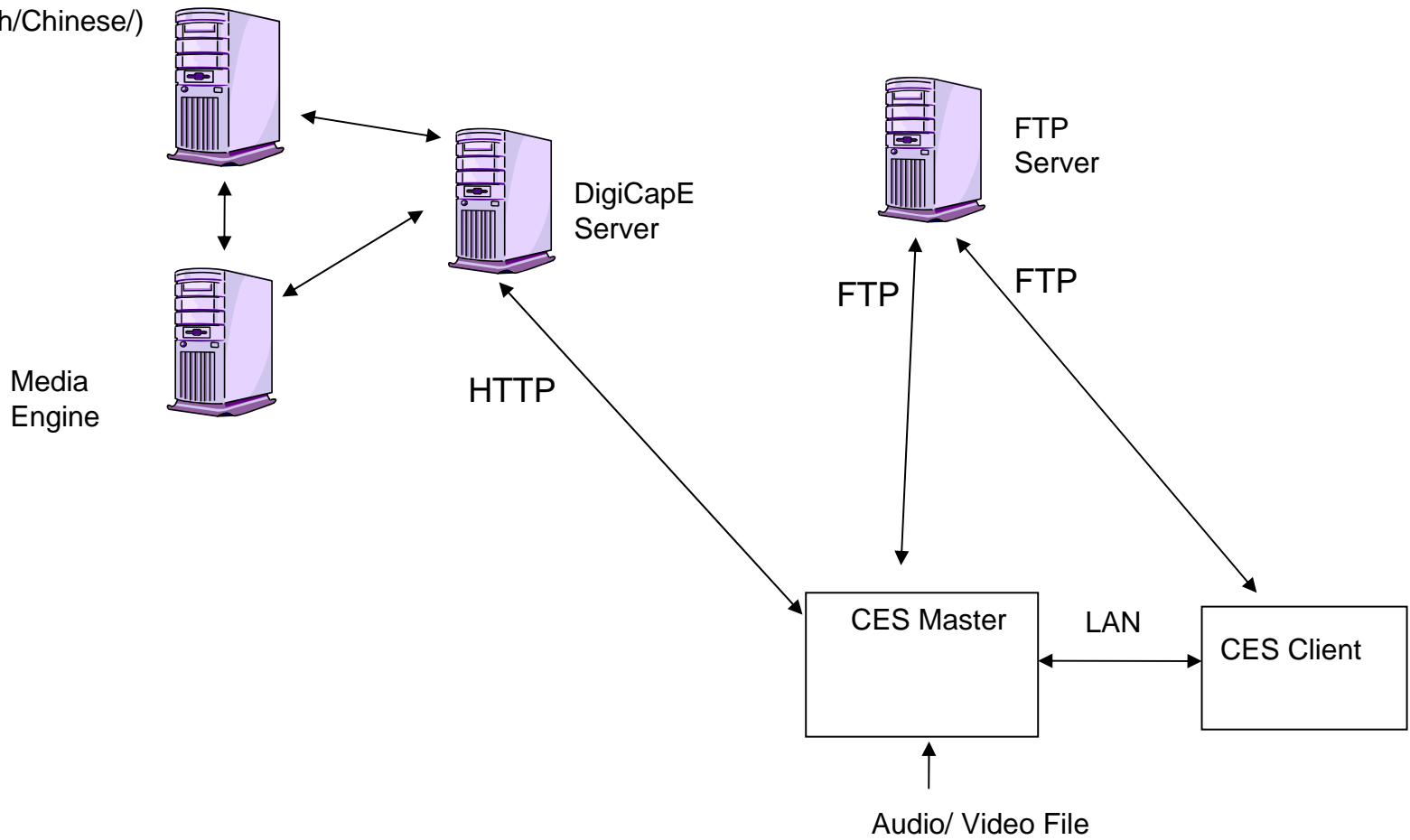
<Key Input>
Insert M, correct A to a
<Fix Complete>
<Jump to next error line>





Components

IBM ASR Server Farm
(English/Chinese/)





Media Coverage

Newspaper

National

- 2005/03/15 Asahi Shinbun
- 2005/03/23 Nikkei Sangyou Shinbun
- 2005/03/30 Kagaku Kougyou Nippo

Local

- 2005/03/23 Shinano Mainichi
- 2005/03/23 Yomiuri Shinbun
- 2005/03/23 Nihon Keizai Shinbun
- 2005/06/07 Nihon Keizai Shinbun
- 2005/06/07 Shinano Mainichi
- 2005/06/07 Toushin Journal
- 2005/06/07 Shinshu Minpo

Magazine Article

- 2005/03/22 Weekly ASCII

TV



National

- 2005/06/06 JNN CS (News Bird)

Local

- 2005/03/22 Nagano Asahi Housou
- 2005/03/22 Shinetu Housou
- 2005/06/06 Ueda Cable TV
- 2005/06/06 NHK Nagano
- 2005/06/06 TV Shinshu
- 2005/06/06 Nagano Asahi Housou
- 2005/06/06 Ueda Cable TV

Papers

- HCI International-Las Vegas 2005
- Rehabilitation Kougaku
- 2005 Education System Information
- CSUN 2006
- Closing the Gap 2006
- IEICE 2007
- HCI International-Beijing 2007

Internet Articles

- 2007/10/26 Yahoo! Japan News
- 2007/10/26 Enterprise Watch
- 2007/10/26 PC Watch



Agenda

IBM Digital Multimedia Captioning and Editing System

Caption Editing System

Demo



Demo

Demonstrations:

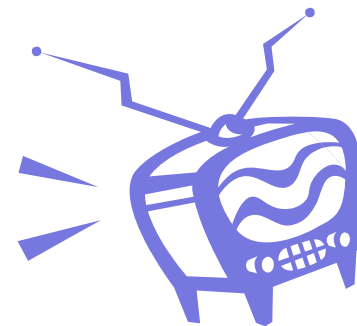
- 1) Transcription by the latest ASR engine
- 2) Edited result of the transcription



Any Questions ??



**Visit us
@
IBM Accessibility Booth # 270
Marriot**





Accessing the Future:

A global collaborative exploration for
accessibility in the next decade

July 20-21, 2009
Boston, MA



Conference tracks:

- Universal Design and Accessibility Standards
- Patient-centered Collaborative Care
- Accessible Online Workplaces and Communities
- Travel & Transportation





The End

Thank You

Contacts:

Ali Sobhi (sobhi@us.ibm.com)

Cole Miyamoto (kmiya@jp.ibm.com)